DATA-DRIVEN STOCHASTIC OPTIMIZATION APPROACHES WITH APPLICATIONS IN POWER SYSTEMS

A Dissertation
Presented to
The Academic Faculty

By

Beste Basciftci

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Operations Research in the
School of Industrial and Systems Engineering

Georgia Institute of Technology

August 2019

DATA-DRIVEN STOCHASTIC OPTIMIZATION APPROACHES WITH APPLICATIONS IN POWER SYSTEMS

Approved by:

Dr. Shabbir Ahmed, Advisor
School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. Nagi Gebraeel, Advisor
School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. Santanu Dey
School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. Andy Sun
School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. Siqian Shen
Industrial & Operations Engineering
University of Michigan

Date Approved: July 9, 2019

To the memory of my beloved grandmother Bahriye Akartuna İğdi,

to my dear mom Aysın Başçiftci and my dear dad Naci Başçiftci,

and to my dear grandfather Besalet İğdi.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere appreciation to my advisors Dr. Shabbir Ahmed and Dr. Nagi Gebraeel, for their invaluable guidance, support and intellectual insights throughout my PhD journey. I was extremely privileged to have the opportunity to work with them, and I would like to thank them in helping me grow academically and become an independent researcher within this journey.

I would like to specifically express my gratitude to Dr. Shabbir Ahmed in his memoriam as he played an instrumental role in my life to study at Georgia Tech. As a masters student in industrial engineering from Turkey, I had been already fascinated by the world of stochastic programming. When I was applying to the PhD programs in US, coming to Georgia Tech and working with Dr. Ahmed were among my dreams. When I was notified to get accepted to ISyE in February 2015, I was very excited as Dr. Ahmed contacted me with a very humble and friendly e-mail, introducing himself and sharing possible research opportunities in his group. Although I was lucky to get accepted from several top programs in our field, working with Dr. Ahmed was a fascinating opportunity for me, and his e-mails changed my life to convince myself to start my PhD at Georgia Tech. I have had the great honor and privilege to work with him since my first day at Georgia Tech, and I am indebted to him for his invaluable guidance and support. He had been a role model for me in every aspect, both academically and personally. As his academic daughter, I miss him dearly... Nevertheless, I would like to emphasize that his legacy will be carried on to the next generations, and his brilliant research ideas, preeminent leadership in our field and his very kind personality will be remembered forever.

I would like to sincerely thank Dr. Santanu Dey, Dr. Siqian Shen and Dr. Andy Sun for serving in my thesis committee, and providing insightful comments. I am extremely grateful for their support during my academic journey.

I was very fortunate to study at ISyE, that provided me such an exciting and active academic environment flourishing my academic and personal pursuits. I also feel grateful to learn many courses from the pioneers of their fields. I would like to specifically express my appreciation to Dr. Edwin Romeijn, Dr. Alan Erera and Dr. Dawn Strickland for their support throughout my years at ISyE. Additionally, I have been in the organization committee of the Discrete Optimization Seminars since the early years of my PhD. It has been an invaluable experience for me, and I would like to thank my friends with whom we organize these seminar series together.

I feel deeply grateful to be surrounded by great friends in ISyE, who have become my Atlanta family. We have many memories that I will forever treasure. I would like to specifically thank them for all their support, encouragement and friendship in the past four years. Additionally, I would like to thank all of my friends from Georgia Tech Turkish Student Organization that helped me to feel at home in Atlanta.

I would like to specifically thank dear Burak, who has been my pillar of strength in these four years during many ups and downs despite of the thousands of miles between us. This thesis would not have been possible without his support, encouragement and motivation throughout my PhD journey.

Finally, I would like to express my deepest appreciation to my dear family for all their love, encouragement and patience. They have been always there for me with all their support and affection. I would like to specifically express my heartfelt appreciation to my grandmother Bahriye and my mother Aysın for their love, affection and support, and helping with me to cultivate my inner strength. There are not enough words to express my gratitude to them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

In this thesis, we focus on data-driven stochastic optimization problems with an emphasis in power systems applications. On the one hand, we address the inefficiencies in maintenance and operations scheduling problems which emerge due to disregarding the uncertainties, and not utilizing statistical analysis methods. On the other hand, we develop a partially adaptive general purpose stochastic programming approach for effectively modeling and solving a class of problems in sequential decision-making.

In the first part of the thesis (Chapter 2 and Chapter 3), we consider maintenance and operations scheduling problem in power systems under uncertainty by leveraging data analytics. In Chapter 2, we develop a stochastic optimization framework for the integrated condition-based maintenance and operations scheduling problem with explicit consideration of sensor-driven unexpected failures. Our approach is based on a model that uses condition-based failure scenarios derived from the remaining lifetime distributions of the generators, as well as a chance constraint to ensure a reliable maintenance plan. The large number of failure scenarios are handled by a combination of sample average approximation and an enhanced scenario decomposition algorithm in a distributed framework. We introduce a number of algorithmic improvements by exploiting the polyhedral structure of the problem, utilizing its time decomposability, and an analysis of the transmission line capacities. Finally, we present a case study demonstrating the significant cost savings and computational benefits of the proposed framework. In Chapter 3, we focus on the tight coupling between the condition of the generators and corresponding operational schedules, significantly affecting reliability of the system. We effectively model and solve an integrated condition-based maintenance and operations scheduling problem for a fleet of generators with an explicit consideration of decision-dependent generator conditions. We propose a sensor-driven degradation framework with remaining lifetime estimation procedures under time varying load levels. We present estimation methods by adapting our model to the un-

derlying signal variability. Then, we develop a stochastic optimization model that considers the effect of the operational decisions on the generators' degradation levels along with the uncertainty of the unexpected failures. As the resulting problem includes nonlinearities, we adopt piecewise linearization along with other linearization techniques and propose formulation enhancements to obtain a stochastic mixed-integer linear programming formulation. We develop a decision-dependent simulation framework for assessing the performance of a given solution. Finally, we present computational experiments demonstrating significant cost savings and reductions in failures in addition to highlighting computational benefits of the proposed approach.

In the second part of the thesis (Chapter 4), we focus on developing a new adaptive stochastic optimization approach for optimizing sequential decision-making processes under uncertainty, which is an inherently complex problem. Two-stage and multi-stage stochastic programming are fundamental techniques for modeling these processes, where stage refers to the decision times in planning. Although both approaches have their individual benefits and limitations, the resulting policies may not be sufficient to address a wide range of business settings due to the level of flexibility required in these processes. To address these settings and find a compromising solution, we propose a novel stochastic programming methodology, labeled as *adaptive two-stage stochastic programming*, in which the stage times are not predetermined but part of the optimization problem. We provide a generic formulation for the proposed approach under finite stochastic processes and prove that the problem is NP-Hard. We also demonstrate the value of this approach by deriving analytical bounds compared to two-stage and multi-stage stochastic programming methods on a special structure that encompasses various problems including capacity expansion planning. To solve the resulting problem, we develop algorithms with approximation guarantee. Our computational studies on sample generation expansion planning instances highlight the importance of the choice of the revision decisions and demonstrate the benefits of adopting the proposed approach from various perspectives.

CHAPTER 1

INTRODUCTION AND BACKGROUND

1.1 Overview of Stochastic Programming

Decision making under uncertainty is a crucial issue as most of the real-life problems involve ambiguities. There are several methods to address this issue depending on the level of knowledge regarding the underlying ambiguity and the restrictions of the specific problem of concern. In particular, stochastic programming plays a key role in this regard to model complex problems, once the distribution of the underlying uncertainty is known.

Let us first consider a generic problem as follows:

$$\min_{x} \quad F(x, \xi) \tag{1.1a}$$

$$\text{st.} \quad G(x, \xi) \leq 0, \tag{1.1b}$$

where the variable x represents our decisions, and the parameter ξ represents data. A deterministic program turns into a stochastic program when the parameter ξ is uncertain. One way to handle the uncertainty in the objective function is to convert it into $\mathbb{E}[F(x, \xi)]$. This is referred as a risk-neutral approach by replacing the objective function with its expected value. As an alternative, one may prefer risk-averse approaches by adding a risk assessment measure to the objective function.

Accurately representing uncertainty is an important concern in stochastic programming. To approximate the underlying uncertainty, sampling is a widely preferred method. The generic problem (1.1) can be approximated as in (1.2) through sampling with scenarios,

$$\min_{x} \quad \frac{1}{K}\sum_{k=1}^{K} F(x,\xi^k) \tag{1.2a}$$

$$\text{st.} \quad G(x,\xi^k) \leq 0 \quad k = 1,\cdots,K, \tag{1.2b}$$

where ξ^k for $k = 1,\cdots,K$ represents the specific realizations of the uncertain parameter ξ. Here, the critical question becomes identifying the necessary number of scenarios for representing the underlying uncertainty accurately.

In several problem contexts, it is essential to obtain decisions that satisfy certain restrictions with a high probability such as ensuring the quality of service in different application areas. To represent these restrictions formally, a *chance-constraint* can be defined as

$$\text{Pr}(G(x,\xi) \leq 0) \geq 1 - \epsilon,$$

where $\epsilon \in (0,1)$. This constraint ensures that the restriction $G(x,\xi) \leq 0$ is satisfied with at least $1-\epsilon$ probability. As chance-constraints are generally computationally intractable, one can consider alternative ways for its representation. The most common methods are formulating its approximations and providing alternative formulations through sampling.

Another important dimension of stochastic programming is related with the timing of the decisions, especially when the planning involves a multi-period problem. On the one hand, the decision-maker can determine a set of decisions ahead of the planning and adjust the remainder ones based on specific realizations. This approach is referred as *two-stage stochastic programming*. On the other hand, the decision-maker can adjust all of her decisions throughout the planning horizon as uncertainty is revealed. This method is considered as *multi-stage stochastic programming* and involves a more complex modeling procedure. Since these problems result in large-scale models as the uncertainty is approximated through scenarios, it is crucial to determine the correct modeling and solution approach to address problems that consider uncertainty over a multi-period planning horizon.

1.2 Overview of Power Systems

Power systems is concerned with the generation, transmission, and distribution of power over networks with its economic aspects. Optimization of power systems can be viewed as a hierarchical planning problem, which can be categorized into three groups with respect to the length of the planning horizon. Long-term planning problems focus on strategic level problems such as generation and transmission capacity expansion planning in order to satisfy future demand forecasts. Medium-term plans focus on monthly or yearly plans. Asset management of key equipments in power systems is an example for this group which involves determining the maintenance schedule of the generators and transmission lines. In short-term planning, operations planning such as unit commitment and economic dispatch are mainly considered.

Power systems operate under the premise of uncompromising generation capacity to satisfy the operational requirements of the network. Generator maintenance scheduling plays a pivotal role in this regard as it determines both the availability and the reliability of the generation resources. Traditionally, vertically integrated utility companies have used different strategies to mitigate the effects of unexpected failures ranging from combinations of supplemental reserves to periodic maintenance intervals. Increased congestion coupled with an aging power grid infrastructure have generated new challenges that cannot be efficiently addressed by the existing strategies. Recently, there has been a growing trend aimed at leveraging sensors to monitor physical and performance degradation of capital-intensive power generating assets through a process commonly referred to as "condition monitoring". Ideally, the goal is to utilize data generated by condition monitoring to predict the remaining operational life of the generator, and use that information to mitigate the risks of unexpected failures. Consequently, it becomes critical to incorporate sensor information to the subsequent decision making processes. In this thesis, we focus on the value of the stochastic optimization through the applications of power systems while leveraging data analytics.

1.3 Organization

In the first part of this thesis, we focus on data-driven stochastic optimization problems for the sensor-driven maintenance and operations scheduling problems in power systems. In Chapter 2, we address the integrated condition-based generator maintenance scheduling problem under the possibility of unexpected failures. This chapter is based on the paper [1]. In Chapter 3, we improve the reliability of the power systems by explicitly considering the tight coupling between generator conditions and their production schedules while optimizing their maintenance and operations. This chapter is based on the paper [2].

In the second part of the thesis, i.e. Chapter 4, we focus on developing a novel adaptive stochastic optimization approach for improving sequential decision making processes under uncertainty with partial flexibility. This chapter presents a case study for the generation expansion planning problem and is based on the paper [3].

CHAPTER 2

STOCHASTIC OPTIMIZATION OF MAINTENANCE AND OPERATIONS SCHEDULES UNDER UNEXPECTED FAILURES

2.1 Introduction

Degradation-based predictive modeling of the generator conditions demonstrated significant cost savings in determining their maintenance schedules (see recent studies [4, 5, 6]). However, the predictive model was assumed to be very accurate, and the impact of unexpected failures on operations was not accounted in maintenance scheduling. In this chapter, we propose a comprehensive framework that uses condition monitoring data to identify optimal maintenance and operational decisions while also modeling the impact of generator failure scenarios.

Scheduling generator maintenance has been studied extensively in the power systems literature [7]. Research on this topic can be broadly classified into two groups. The first group revolves around maintenance optimization ([8, 9, 4]), while the second integrates maintenance with operational decisions ([10, 11, 12, 13, 14, 15, 5, 6]). Most of these models determine maintenance schedules regardless of the generator's condition. The few examples that utilize condition monitoring assume that failures are completely predictable. In contrast, our approach is based on a stochastic optimization framework that accounts for unexpected generator failures, yet still leverages condition monitoring information for the joint optimization of maintenance and operations scheduling. Stochastic optimization has been widely used for optimizing operational problems, see [16] for a recent review. Some studies ([17, 18, 19]) consider uncertainty at the operational level of the joint optimization problem by taking into account a limited number of scenarios with demand and price uncertainty. However, there has been little to no emphasis on failure uncertainty, much less

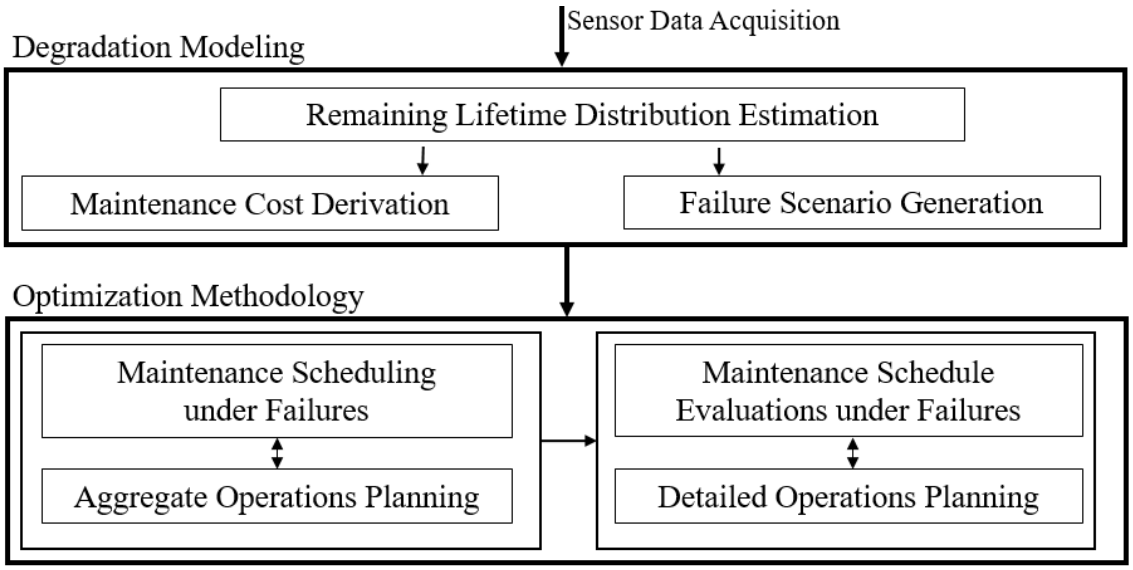


Figure 2.1: Flowchart of the Proposed Framework

on joint maintenance-operations planning under uncertainty.

In this chapter, we develop and solve a joint maintenance and operations planning problem with explicit consideration of unexpected generator failures. The proposed approach is relevant to vertically integrated utilities that have the authority to determine maintenance schedules for its entire generation fleet. Figure 2.1 presents an overview of the proposed framework. We first model the condition of the generators using a sensor-based degradation modeling framework, as the one proposed in [20]. Unique to our methodology, we use predicted remaining life distributions to construct scenarios of generator failures, and integrate them in maintenance scheduling. Furthermore, we propose a multi-scale methodology by first performing joint optimization of daily maintenance and operations schedules, and then evaluating the resulting maintenance schedules over a more detailed "hourly" operational problem under unexpected failure scenarios.

The contributions of this chapter are summarized as follows:

1. We develop a stochastic optimization framework for modeling sensor-driven maintenance and operations schedules under unexpected generator failures. Specifically, we introduce a stochastic programming model for jointly optimizing maintenance

and operations schedules with reliability and cost perspectives. The proposed model differs from existing studies in that:

   (a) Failure scenarios are generated using degradation-based remaining life distributions of the generators and are explicitly considered in the joint optimization problem.

   (b) We introduce a chance constraint to restrict the number of generator failures, and develop a deterministic safe approximation.

2. We propose a solution methodology by enhancing a generic scenario decomposition approach [21] with a number of algorithmic improvements specific to our problem. These enhancements include, a) generating stronger cuts, b) utilizing the time decomposability of the algorithm to minimize the number of re-solves, and c) leveraging parallel computing for algorithm implementation.

3. We demonstrate the effectiveness of the proposed approach on three IEEE test cases [22, 23, 24]. Our solution algorithm exhibits an almost ideal parallel speed up and significant computational gains relative to its generic version due to the introduced enhancements. The sampling scheme allows us to obtain maintenance solutions within 2% optimality gap for problems with up to 13^{19} scenarios. Moreover, the computational experiments highlight 7-19% cost savings using the proposed stochastic programming approach relative to deterministic schedules.

The chapter is organized as follows: Section 2.2 discusses the integrated condition-based generator maintenance and operations planning problem. It presents an overview of the degradation modeling framework, the scenario generation procedure, and the joint optimization model. Section 2.3 introduces the scenario decomposition algorithm along with our methodological enhancements. Section 2.4 presents the computational experiments and results. Section 2.5 provides the concluding remarks of this chapter.

2.2 Model Formulation

2.2.1 Degradation modeling and scenario generation

Degradation is defined as the progressive accumulation of damage due to natural wear and tear. Generators like many mechanical equipment and machinery degrade over time. In most cases, physical degradation can be monitored using sensors through a process known as "condition monitoring". In this work, we assume that the degradation-based sensor signal of generator i can be modeled as a continuous stochastic function $S_i(t)$. Furthermore, we assume that the signal increases in severity until it exceeds a prespecified alarm threshold, Λ_i. A stochastic degradation model proposed in [4] is then used to predict the statistical distribution of the remaining lifetime of the generator, hereafter referred to as the RLD of the generator. The RLDs of the generators can be updated, in real-time using a Bayesian framework as more sensor data is observed. We used the RLD estimation procedure proposed in [20] and [25], which solely focuses on stochastic degradation modeling of components. This procedure is subsequently integrated into a deterministic maintenance scheduling problem in [4] and [5]. However, these studies make a key assumption that the generators would not fail under sensor-driven maintenance schedules. Thus, one of our main contributions is to propose a stochastic optimization framework to model generator maintenance and operations scheduling problem under the uncertainty of the generator failures.

Different to our methodology, the RLDs of the generators are used to generate failure scenarios. A scenario consists of possible failure times for each generator. To do this, we use the most recently updated RLDs of the generators to determine the probability π_k associated with each failure scenario k. To arrive at a finite set of scenarios, a planning horizon of length H is discretized into d time intervals each corresponding to a possible failure period. An additional period is added to account for generator non-failures, i.e., a generator does not fail within the planning horizon. The probability that generator i

fails in the range j is calculated for every generator i and range $j = 1, \cdots, d$, using RLD of each generator i, namely F_i. The probability that generator i does not fail within the planning horizon is given by Pr$(R_i^{t_i^o} > H)$, where $R_i^{t_i^o}$ is the residual life of generator i at the observation time t_i^o. We denote the failure time of generator i in scenario k by τ_i^k, and assume that failure occurs in the middle of the selected failure period. For non-failures, τ_i^k is set to $H + 1$.

One can potentially consider all failure cases by extensively generating all scenarios. However, this becomes computationally challenging as the number of scenarios grows exponentially in the extensive case, resulting in $(d + 1)^{|\mathcal{G}|}$ many scenarios in total. In order to overcome this issue, we propose an alternative scenario generation procedure based on sampling. In this procedure, we generate independently and identically distributed samples, which is described in Algorithm 1.

Algorithm 1 Scenario generation with SN samples and d discretizations

Set $T_0 = 0$, $T_{d+1} = \infty$, $T_j = \left\lfloor j\dfrac{H}{d} \right\rfloor$, $j = 1, \ldots, d$.
Obtain the discretized ranges in the planning horizon: $[T_{j-1}, T_j)$ for $j = 1, \ldots, d + 1$.
for all $k = 1, \ldots, SN$ **do**
    for all $i = 1, \ldots, |\mathcal{G}|$ **do**
        Generate U from Uniform(0,1).
        Find j s.t. $F_i(T_{j-1}) \leq U < F_i(T_j)$.
        if $j \leq d$ **then**
            $\tau_i^k = \lfloor (T_{j-1} + T_j)/2 \rfloor$.
        else
            $\tau_i^k = H + 1$.
        end if
    end for
    Set $\pi_k = 1/SN$.
end for

2.2.2 Optimization model

We formulate the integrated generator maintenance and operations scheduling problem as a stochastic mixed-integer program that considers unexpected generator failure scenarios.

The model aims to minimize the expected maintenance and operational costs while determining optimal maintenance schedules and operational decisions specific to each scenario. We consider a finite planning horizon that consists of maintenance periods with operational subperiods for commitment decisions, dispatch and demand curtailment amounts. We assume that each generator experiences a single maintenance routine, preventive or corrective, within the planning horizon. Generators cannot produce electricity during maintenance. A corrective maintenance (CM) is performed if a generator fails unexpectedly before its predicted maintenance period with a duration of Y_c periods and a cost, C_c. Otherwise, predictive maintenance (PM) is conducted at the scheduled maintenance period, which lasts Y_p periods and costs $C_{i,t}$. Here, we note that $Y_c > Y_p$ since CM is conducted under emergency conditions that typically require unplanned dispatch of maintenance resources. The term $C_{i,t}$ is the PM cost function, i.e., a function that determines the cost of performing a PM at different time periods. The PM cost function is calculated using equation (2.1) [4], which uses the generator's most recent RLD prediction to balance the costs of early maintenance and unexpected failure, and is represented as follows:

$$C_{i,t}^{t_i^o} = \frac{c_i^p \, \Pr(R_i^{t_i^o} > t) + c_i^c \, \Pr(R_i^{t_i^o} \le t)}{\int_0^t \Pr(R_i^{t_i^o} > z)dz \ + t_i^o}, \tag{2.1}$$

where $C_{i,t}^{t_i^o}$ is the PM cost of generator i after t maintenance epochs from the observation time t_i^o, c_i^p is the cost of early maintenance and c_i^c is the cost of unexpected failure. The observation time t_i^o is taken as the time of planning for all generators, and hence omitted in the formulation. Sets, decision variables, and parameters of the optimization model can be summarized as follows:

Sets:

$\mathcal{B}$ Set of buses.

$\mathcal{G}$ Set of generators.

$\mathcal{K}$ Set of scenarios.

$\mathcal{L}$ Set of transmission lines.

$\mathcal{S}$ Set of operational periods in a maintenance period.

$\mathcal{T}$ Set of maintenance periods in the planning horizon.

Decision Variables:

$z_{i,t}$ $z_{i,t} = 1$ if generator i enters maintenance in maintenance period t, and 0 otherwise.

γ_t^k Additional maintenance capacity added in maintenance period t.

$\eta_{i,t,s}^k$ $\eta_{i,t,s}^k = 1$ if generator i is on in operational period s of maintenance period t in scenario k, and 0 otherwise.

$w_{i,t,s}^k$ Dispatch amount of generator i in operational period s of maintenance period t in scenario k.

$u_{i,t,s}^k$ $u_{i,t,s}^k = 1$ if generator i starts up in operational period s of maintenance period t in scenario k, and 0 otherwise.

$v_{i,t,s}^k$ $v_{i,t,s}^k = 1$ if generator i shuts down in operational period s of maintenance period t in scenario k, and 0 otherwise.

$\psi_{b,t,s}^k$ Demand curtailed in bus b in operational period s of maintenance period t in scenario k.

Parameters:

C_{add} Per unit cost of maintenance capacity added.

C_c Corrective maintenance cost.

$C_{i,t}$ Predictive maintenance cost of generator i in maintenance period t.

$F_{i,t,s}$ Per unit dispatch cost of generator i in operational period s of maintenance period t.

H Planning horizon length in maintenance periods.

L Capacity on the ongoing number of maintenances.

$N_{i,t,s}$ No-load cost of generator i in the operational period s of maintenance period t.

P_{DC} Per unit cost of demand curtailed.

$U_{i,t,s}$ Start-up cost of generator i in operational period s of maintenance period t.

$V_{i,t,s}$ Shut-down cost of generator i in operational period s of maintenance period t.

τ_i^k Failure time of generator i in scenario k.

π_k Probability of scenario k.

ϵ Confidence level of the chance constraint.

ρ Threshold on the number of generators to fail.

$D_{b,t,s}$ Demand of bus b in operational period s of maintenance period t.

f_{max}^l Flow capacity of line l.

a_l Shift factor vector for line l.

$M_{b,i}$ Generation location matrix, which is 1 if generator i is on bus b, and 0 otherwise.

p_i^{min} Minimum production requirement of generator i.

p_i^{max} Maximum production capacity of generator i.

RD_i Ramp-down rate of generator i.

RU_i Ramp-down rate of generator i.

The proposed optimization model is formulated as follows:

$$\min \sum_{k\in\mathcal{K}} \pi_k \left(\sum_{i\in\mathcal{G}} \sum_{t=1}^{\tau_i^k-1} C_{i,t} z_{i,t} + \sum_{i\in\mathcal{G}} \sum_{t=\tau_i^k}^{H} C_c z_{i,t} \right) + \sum_{k\in\mathcal{K}} \sum_{i\in\mathcal{G}} \sum_{t\in\mathcal{T}} \sum_{s\in\mathcal{S}} \pi_k \left(N_{i,t,s} \eta_{i,t,s}^k + F_{i,t,s} w_{i,t,s}^k \right)$$
$$+ \sum_{k\in\mathcal{K}} \sum_{i\in\mathcal{G}} \sum_{t\in\mathcal{T}} \sum_{s\in\mathcal{S}} \pi_k \left(U_{i,t,s} u_{i,t,s}^k + V_{i,t,s} v_{i,t,s}^k \right) + \sum_{k\in\mathcal{K}} \sum_{t\in\mathcal{T}} \sum_{s\in\mathcal{S}} \sum_{b\in\mathcal{B}} \pi_k P_{DC} \psi_{b,t,s}^k$$
$$+ \sum_{k\in\mathcal{K}} \sum_{t\in\mathcal{T}} \pi_k C_{add} \gamma_t^k \tag{2.2a}$$

$$\text{s.t. } \Pr\left(\sum_{i\in\mathcal{G}} \sum_{t\in\mathcal{T}} \zeta_{i,t} z_{i,t} \le \rho \right) \ge 1-\epsilon \tag{2.2b}$$

$$\sum_{\substack{i\in\mathcal{G}: \\ t\le\tau_i^k-1}} \sum_{e=0}^{Y_p-1} z_{i,t-e} \le L + \gamma_t^k \qquad t\in\mathcal{T}, k\in\mathcal{K} \tag{2.2c}$$

$$\sum_{t\in\mathcal{T}} z_{i,t} = 1 \qquad i\in\mathcal{G} \tag{2.2d}$$

$$\eta_{i,t,s}^k \le 1 - \sum_{e=0}^{Y_p-1} z_{i,t-e} \qquad i\in\mathcal{G}, t\in\{1,\ldots,\tau_i^k+Y_p-1\}, s\in\mathcal{S}, k\in\mathcal{K} \tag{2.2e}$$

$$\eta_{i,t,s}^k \le \sum_{t'=1}^{\tau_i^k-1} z_{i,t'} \qquad i\in\mathcal{G}, t\in\{\tau_i^k,\ldots,\tau_i^k+Y_c-1\}, s\in\mathcal{S}, k\in\mathcal{K} \tag{2.2f}$$

$$\eta_{i,t,s-1}^k - \eta_{i,t,s}^k + u_{i,t,s}^k \ge 0 \qquad i\in\mathcal{G}, t\in\mathcal{T}, s\in\mathcal{S}, k\in\mathcal{K} \tag{2.2g}$$

$$\eta_{i,t,s}^k - \eta_{i,t,s-1}^k + v_{i,t,s}^k \ge 0 \qquad i\in\mathcal{G}, t\in\mathcal{T}, s\in\mathcal{S}, k\in\mathcal{K} \tag{2.2h}$$

$$\sum_{i\in\mathcal{G}} w_{i,t,s}^k + \sum_{b\in\mathcal{B}} \psi_{b,t,s}^k = \sum_{b\in\mathcal{B}} D_{b,t,s} \qquad t\in\mathcal{T}, s\in\mathcal{S}, k\in\mathcal{K} \tag{2.2i}$$

$$\left| \sum_{b\in\mathcal{B}} a_{l,b} \left(\sum_{i\in\mathcal{G}} M_{b,i} w_{i,t,s}^k + \psi_{b,t,s}^k - D_{b,t,s} \right) \right| \le f_l^{max} \qquad l\in\mathcal{L}, t\in\mathcal{T}, s\in\mathcal{S}, k\in\mathcal{K} \tag{2.2j}$$

$$p_i^{min} \eta_{i,t,s}^k \le w_{i,t,s}^k \le p_i^{max} \eta_{i,t,s}^k \qquad i\in\mathcal{G}, t\in\mathcal{T}, s\in\mathcal{S}, k\in\mathcal{K} \tag{2.2k}$$

$$-RD_i \le w_{i,t,s}^k - w_{i,t,s-1}^k \le RU_i \qquad i\in\mathcal{G}, t\in\mathcal{T}, s\in\mathcal{S}, k\in\mathcal{K} \tag{2.2l}$$

$$z_{i,t}, \eta_{i,t,s}^k, u_{i,t,s}^k, v_{i,t,s}^k \in \{0,1\}, \gamma_t^k, w_{i,t,s}^k, \psi_{b,t,s}^k \ge 0 \quad i\in\mathcal{G}, t\in\mathcal{T}, s\in\mathcal{S}, k\in\mathcal{K}, b\in\mathcal{B}. \tag{2.2m}$$

The objective function (2.2a) consists of five parts. The first part corresponds to the expected maintenance cost. If the maintenance of generator i is scheduled before its failure time in scenario k, then a PM cost is incurred, otherwise the cost of CM is incurred. The second, third and fourth parts of the objective function represent the expected operational cost and accounts for commitment, dispatch, demand curtailment, start up, and shut down costs. The last part corresponds to the expected cost of additional PMs. Additional PMs are penalized with the cost, C_{add} that accounts for the additional maintenance resources required for performing multiple repair tasks simultaneously.

Constraint (2.2b) is a chance constraint which ensures that the number of generators under CM is less than a predefined threshold ρ with probability $1-\epsilon$. This constraint fully adapts to the sensor information since $\zeta_{i,t}$ is a random variable that takes on a value of 1 if $t \geq \tau_i$ and 0 otherwise. The distribution of τ_i is determined by the estimated RLD of generator i. As this constraint is intractable, safe approximations are derived in Section 2.2.2. Constraint (2.2c) ensures that there can be at most $L + lb_t^k$ ongoing PMs for each period t and scenario k. Constraint (2.2d) ensures that one maintenance must be scheduled within the planning horizon for every generator. Constraints (2.2e) and (2.2f) impose the logical relationship between a generator's failure at a specific scenario and the maintenance decision. That is, for a given scenario k, if the planned maintenance $z_{i,t}$ is before the generator's failure time τ_i^k, then constraint (2.2e) guarantees that generator i is under PM, and is unavailable during the interval $[z_{i,t}, z_{i,t} + Y_p - 1]$. Otherwise, constraint (2.2f) guarantees that the generator undergoes a CM, and is unavailable during the interval $[\tau_i^k, \tau_i^k + Y_c - 1]$. Constraints (2.2g) and (2.2h) are used for modeling the coupling between commitment, start-up and shut-down variables. Constraint (2.2i) ensures that the total demand is satisfied. Constraints (2.2j) and (2.2k) satisfy transmission line capacity restrictions, and generator production capacity restrictions, respectively. Constraint (2.2l) guarantees that the production difference is between ramp up and ramp down limits.

Our main focus is to analyze the effect of unexpected generator failures on maintenance and operations scheduling. Thus, the only source of uncertainty in the optimization model is the randomness of the failure times, and deterministic demand values are considered. We propose a stochastic optimization approach, which can be interpreted as a combination of a stochastic programming and a robust optimization to handle the uncertainty. Adopting solely a robust optimization approach might result in conservative maintenance schedules by considering the worst-case scenario. Thus, the robust solutions might enforce unnecessarily early preventive maintenances on the entire fleet of generators. For handling this issue, we put forward a more balanced approach by taking into account various failure scenarios and considering the expected maintenance and operational costs through stochastic programming. In order to ensure that most of the generators enter PM, we propose a chance-constraint that restricts the probability of the generators that enter CM. The deterministic safe approximations of the chance-constraint, discussed in the next section, can still be interpreted as a robust approach, as we conservatively restrict the number of generators that enter maintenance due to failure.

Safe approximations of the chance constraint (2.2b)

Chance constraint (2.2b) can be reexpressed as follows; $\Pr\left(\sum_{i\in\mathcal{G}}\sum_{t\in\mathcal{T}}\zeta_{i,t}z_{i,t}\geq\rho\right)\leq\epsilon$. A safe approximation can be found by computing an upper bound on the probability expression on left hand side of this inequality. To do this, we use Markov and generalized Bernstein inequalities (see [26]).

Proposition 1. *The deterministic linear constraint (2.3) is a safe approximation of* (2.2b)*, i.e. any $z\in\{0,1\}^{|\mathcal{T}|\times|\mathcal{G}|}$ satisfying (2.3), satisfies* (2.2b).

$$\sum_{i\in\mathcal{G}}\sum_{t\in\mathcal{T}}\mathbf{E}[\zeta_{i,t}]\,z_{i,t}\leq\max\left(\rho\,\epsilon,\max_{\alpha>0}\left[\frac{((\epsilon\,e^{\alpha\rho})^{1/|\mathcal{G}|}-1)|\mathcal{G}|}{e^{\alpha}-1}\right]\right)\tag{2.3}$$

Proof. We prove the result in two parts. First, using Markov's inequality and linearity of expectation, we have:

$$\begin{aligned}\Pr(\sum_{i\in\mathcal{G}}\sum_{t\in\mathcal{T}}\zeta_{i,t}z_{i,t}\geq\rho) &\leq \mathbf{E}[\sum_{i\in\mathcal{G}}\sum_{t\in\mathcal{T}}\zeta_{i,t}z_{i,t}]/\rho\\ &= \sum_{i\in\mathcal{G}}\sum_{t\in\mathcal{T}}\mathbf{E}[\zeta_{i,t}]z_{i,t}/\rho.\end{aligned}$$

Hence, $\sum_{i\in\mathcal{G}}\sum_{t\in\mathcal{T}}\mathbf{E}[\zeta_{i,t}]\ z_{i,t}/\rho\leq\epsilon$ provides a safe approximation of the constraint (2.2b).

Second, let $G_i(z)=\sum_{t\in\mathcal{T}}\zeta_{i,t}z_{i,t}$ be a random variable depending on generator i and a maintenance decision z. From constraint (2.2d), we see that $\sum_{t\in\mathcal{T}}z_{i,t}=1$, thus, $G_i(z)$ is a Bernoulli random variable that is equal to 1 with probability $p_i(z)$, and 0 otherwise. Here, $p_i(z)=\sum_{t\in\mathcal{T}}\Pr(t\geq\tau_i)\ z_{i,t}$ represents the probability that generator i fails under the solution z. Thus, $\Pr\left(\sum_{i\in\mathcal{G}}\sum_{t\in\mathcal{T}}\zeta_{i,t}z_{i,t}\geq\rho\right)$ can be rewritten as $\Pr\left(\sum_{i\in\mathcal{G}}G_i(z)\geq\rho\right)$. Given a maintenance decision z, $G_i(z)$ for each generator i are independently distributed. Thus, for any $\alpha>0$, we have;

$$\begin{aligned}\Pr\left(\sum_{i\in\mathcal{G}}G_i(z)\geq\rho\right) &= \Pr\left(e^{\alpha\sum_{i\in\mathcal{G}}G_i(z)}\geq e^{\alpha\rho}\right)\\ &\leq \frac{\mathbf{E}[e^{\alpha\sum_{i\in\mathcal{G}}G_i(z)}]}{e^{\alpha\rho}}\\ &= \frac{\mathbf{E}[\prod_{i\in\mathcal{G}}e^{\alpha\,G_i(z)}]}{e^{\alpha\rho}}\\ &= \frac{\prod_{i\in\mathcal{G}}\mathbf{E}[e^{\alpha\,G_i(z)}]}{e^{\alpha\rho}}\\ &= \frac{\prod_{i\in\mathcal{G}}[p_i(z)\ e^{\alpha}+(1-p_i(z))]}{e^{\alpha\rho}}\\ &\leq \frac{[p(z)\ e^{\alpha}+(1-p(z))]^{|\mathcal{G}|}}{e^{\alpha\rho}},\end{aligned}$$

where $p(z)=\frac{\sum_{i\in\mathcal{G}}p_i(z)}{|\mathcal{G}|}$.

The first inequality follows from Markov's inequality, and the last inequality follows from the geometric-arithmetic means inequality. By upper bounding the resulting expression by ϵ, we obtain a safe approximation, $\frac{[p(z)\ e^{\alpha}+(1-p(z))]^{|\mathcal{G}|}}{e^{\alpha\rho}}\leq\epsilon$ for the constraint (2.2b).

Substituting the value of $p(z) = \frac{\sum_{i \in \mathcal{G}} p_i(z)}{|\mathcal{G}|} = \frac{\sum_{i \in \mathcal{G}} \sum_{t \in \mathcal{T}} \Pr(t \geq \tau_i)\, z_{i,t}}{|\mathcal{G}|}$, we have

$$\sum_{i \in \mathcal{G}} \sum_{t \in \mathcal{T}} \Pr(t \geq \tau_i)\, z_{i,t} \leq \frac{((\epsilon\, e^{\alpha\rho})^{1/|\mathcal{G}|} - 1)|\mathcal{G}|}{e^{\alpha} - 1}.$$

Since $\mathbf{E}[\zeta_{i,t}] = \Pr(t \geq \tau_i)$, we obtain the desired bound. The proposed inequality provides a safe approximation to the chance-constraint (2.2b) for any α positive. To achieve a least conservative approximation, we select the α value that maximizes the right hand side of the constraint. □

Decomposition Structure

The stochastic program (2.2) can be compactly represented as:

$$\min \sum_{k \in \mathcal{K}} \pi_k(\quad c_k^\top z + p^\top x^k \quad + f^\top \gamma^k + b^\top y^k) \tag{2.4a}$$

$$\text{s.t.} \quad Az \quad \leq l \tag{2.4b}$$

$$B_k z + E x^k \quad + G\gamma^k \quad \leq m \;\; k \in \mathcal{K} \tag{2.4c}$$

$$F x^k \quad + H y^k \quad \leq n \;\; k \in \mathcal{K} \tag{2.4d}$$

$$z \in \{0,1\}^{|\mathcal{T}| \times |\mathcal{G}|},\; x^k \in \{0,1\}^{3|\mathcal{T}| \times |\mathcal{G}| \times |\mathcal{S}|},\; \gamma^k, y^k \geq 0\; k \in \mathcal{K}.$$

Here, constraint (2.4b) refers to the safe approximation of the chance-constraint (2.2b), i.e. (2.3), and maintenance constraint (2.2d). Constraint (2.4c) represent the coupling constraints between maintenance and operations ((2.2c), (2.2e), (2.2f)). Constraint (2.4d) refers to the operational constraints ((2.2g), (2.2h), (2.2i), (2.2j), (2.2k), (2.2l)). The decision variables are denoted in compact form where x^k represents commitment, start-up, and shut down variables, and y^k represents dispatch and demand curtailment decisions in scenario k.

This formulation is a two-stage stochastic program with first-stage variables z corresponding to maintenance decisions and second-stage variables x, y, and γ corresponding to the operational decisions and the additional labor. This formulation can be represented as

$$\min_{z}\{\sum_{k\in\mathcal{K}}\pi_k f_k(z) : Az \leq l, z \in \{0,1\}^{|\mathcal{T}|\times|\mathcal{G}|}\}, \tag{2.5}$$

where

$$\begin{aligned} f_k(z) = c_k^\top z + \min_{x,y,\gamma}\{&p^\top x^k + f^\top \gamma^k + b^\top y^k : \\ &Ex^k + G\gamma^k \leq m - B_k z, Fx^k + Hy^k \leq n, \\ &x^k \in \{0,1\}^{3|\mathcal{T}|\times|\mathcal{G}|\times|\mathcal{S}|}, \gamma^k, y^k \geq 0\}. \end{aligned} \tag{2.6}$$

Given a maintenance decision z, we can observe that the operational decisions corresponding to each maintenance period t, namely $x^{t,k}$ and $y^{t,k}$, are independent as there are no coupling constraints between maintenance periods in the operational problem. Thus, given a maintenance decision z, the scenario subproblems $f_k(z)$ can be decomposed with respect to the maintenance epochs such that $f_k(z) = c_k^\top z + \sum_{t\in\mathcal{T}} f_k^t(z)$, where $f_k^t(z)$ is defined as:

$$\begin{aligned} f_k^t(z) = \min_{x^t,y^t,\gamma^t}\{&{p^t}^\top x^{t,k} + {f^t}^\top \gamma^{t,k} + {b^t}^\top y^{t,k} : \\ &E^t x^{t,k} + G^t\gamma^{t,k} \leq m^t - B_k^t z, F^t x^{t,k} + H^t y^{t,k} \leq n^t, \\ &x^{t,k} \in \{0,1\}^{3|\mathcal{G}|\times|\mathcal{S}|}, \gamma^{t,k}, y^{t,k} \geq 0\}. \end{aligned} \tag{2.7}$$

Using this formulation we can solve smaller subproblems for each maintenance epoch t and scenario k, which results in significant computational advantage over (2.6).

2.3 Solution Methodology

We solve the two-stage stochastic program (2.5) using a sample average approximation approach by optimizing the stochastic program over a random scenario subset K. The resulting problem is still challenging to solve, and hence requires efficient solution techniques. We address this issue by using a scenario decomposition algorithm and proposing various algorithmic enhancements.

In our solution methodology, we propose a two-step, multi-scale approach, as described in Fig. 2.1. In the first step, we solve an aggregate operational problem to determine the maintenance schedule. In the second step, we evaluate the maintenance schedule on operations at an hourly level.

2.3.1 Sample average approximation methodology

The number of scenarios increases exponentially in the number of generators making the stochastic program (2.5) computationally expensive. Instead, we solve (2.5) over an independently and identically distributed (i.i.d.) scenario subset of different generator failure times as described in Section 2.2.1. We then derive confidence intervals of the true optimal value using the guidelines proposed in [27].

Algorithm 2 outlines the key steps used to derive the confidence intervals using the Sample Average Approximation (SAA) method. First, we generate M batches of N scenarios of generator failures. The resulting SAA problems are solved (Algorithm 2, step 4), and the average of their objectives is used to provide a lower bound estimate of the true optimal value. Next, we evaluate the resulting M feasible solutions over N' scenarios where $N' \gg N$. The solution coming from each batch with the smallest upper bound is selected for evaluating the upper bound estimate of the true optimal value. After obtaining the confidence intervals for the lower and upper bounds, a $(1-\alpha)\%$ confidence interval for the true objective is constructed as follows:

$$(\mu_L - t_{\alpha/2,M-1}\sigma_L/\sqrt{M}, \mu_U + z_{\alpha/2}\sigma_U/\sqrt{N'}) \tag{2.8}$$

2.3.2 Scenario decomposition algorithm

To solve model (2.5) over a scenario sample, we introduce Algorithm 3, an enhanced version of a scenario decomposition algorithm proposed in [21]. Algorithm 3 has two com-

Algorithm 2 SAA method

Generate an i.i.d. scenario sample of size N'.
for all $k = 1, \dots, M$ **do**
    Generate an i.i.d. sample of size N.
    Solve $v_N^k = \min\{\frac{1}{N}\sum_{j=1}^{N} f_j(z) : Az \leq l,\, z \in \{0,1\}^{|\mathcal{T}|\times|\mathcal{G}|}\}$ using Algorithm 3 and obtain z_N^k.
    Evaluate the solution z_N^k over N' scenarios: $\mu_{N'}^k = \frac{1}{N'}\sum_{k\in N'}\left(c_k^\top z_N^k + \sum_{t\in\mathcal{T}} f_k^t(z_N^k)\right)$.
end for
Select the best upper bound estimate: $\mu_U = \min\limits_{k\in\{1,\dots,M\}} \mu_{N'}^k$, and the corresponding solution $\bar{z}$.
Compute variance of the true upper bound estimate, σ_U^2 :

$$\sigma_U^2 = \frac{1}{N'-1}\sum_{k=1}^{N'}\left(\left(c_k^\top \bar{z} + \sum_{t\in\mathcal{T}} f_k^t(\bar{z})\right) - \mu_U\right)^2.$$

Construct the $(1-\alpha)$ level confidence interval for the upper bound estimate as $\mu_U \pm z_{\alpha/2}\sigma_U/\sqrt{N'}$.
Compute mean and variance of the true lower bound estimate, μ_L and σ_L^2 as

$$\mu_L = \frac{1}{M}\sum_{k=1}^{M} v_N^k \text{ and } \sigma_L^2 = \frac{1}{M-1}\sum_{k=1}^{M}(v_N^k - \mu_L)^2.$$

Construct the $(1-\alpha)$ level confidence interval for the lower bound estimate as $\mu_L \pm t_{\alpha/2,M-1}\sigma_L/\sqrt{M}$.

ponents, a lower and an upper bounding step. In the lower bounding step, each scenario subproblem is solved separately to obtain lower bounds and candidate feasible first stage maintenance decisions z. In the upper bounding step, these solutions are evaluated under each scenario setting, and the best upper bound of that iteration is found. The algorithm proceeds by using integer cuts to eliminate the solutions that are already explored in order to obtain new candidate solutions in each iteration. This process is repeated until the lower bound is close enough to the upper bound. Since there are finitely many feasible maintenance solutions, the algorithm is guaranteed to converge in finitely many iterations, as proven in [21].

Utilizing the time decomposability of the second stage problem, we propose an algorithmic improvement that identifies generator statuses for each time period t as a binary vector of size $|\mathcal{G}|$ in the upper bounding step. This vector, namely υ_t, shows whether generators are available for production in period t based on their maintenance status in $\hat{z}$ and failure times τ_i^k in scenario k. This information is sufficient to solve the second stage operational problem for each time t. As a preprocessing step (before the upper bounding part), the unique set of generator statuses for each time period t are identified in the set Υ_t. It becomes sufficient to solve the second stage subproblems only for the unique set of generator statuses found in that iteration. This provides significant computational gains by storing the values corresponding to the previously explored generator statuses.

Algorithm 3 Scenario Decomposition

$LB = -\infty, UB = \infty, \mathcal{Z} = \emptyset, z^* = \emptyset$.

Set $\Upsilon_t = \emptyset, \Phi_t = \emptyset$ for all $t \in \mathcal{T}$.

while $UB > LB$ **do**

    Lower Bounding:

    for all $k \in \mathcal{K}$ *(in parallel)* **do**

        $z^k = \mathrm{argmin}\{f_k(z) : Az \leq l, z \in \{0,1\}^{|\mathcal{T}|\times|\mathcal{G}|} \setminus \mathcal{Z}\}$.

    end for

    Compute $LB = \sum_{k\in\mathcal{K}} \pi_k f_k(z^k)$.

    $\widehat{\mathcal{Z}} = \bigcup_{k\in\mathcal{K}}\{z^k\}$, $\mathcal{Z} = \mathcal{Z} \cup \widehat{\mathcal{Z}}$ {Add the set of explored solutions in this iteration, namely $\widehat{\mathcal{Z}}$, to the set $\mathcal{Z}$.}

    Upper Bounding:

    Identify generator statuses υ_t, over every pair $(\hat{z}, k) \in \widehat{\mathcal{Z}} \times \mathcal{K}$ for all $t \in \mathcal{T}$, and set $\Phi_t = \bigcup_{(\hat{z},k,t)\in\widehat{\mathcal{Z}}\times\mathcal{K}\times\mathcal{T}}\{\upsilon_t, (\hat{z}, k, t)\}$. {Store solution, scenario, time triple $(\hat{z}, k, t)$ with its status υ_t.}.

    Set $\widehat{\Upsilon_t} = \left(\bigcup_{t\in\mathcal{T}}\{\upsilon_t\}\right) \setminus \Upsilon_t$, $\Upsilon_t = \Upsilon_t \cup \widehat{\Upsilon_t}$ for all $t \in \mathcal{T}$.

    for all $t \in \mathcal{T}$ **do**

        for all $\upsilon_t \in \widehat{\Upsilon_t}$ *(in parallel)* **do**

            Obtain a $(\hat{z}, k, t)$ triple that υ_t maps in the set Φ_t.

            Solve (2.7) to obtain $f_k^t(\hat{z})$ and set $f_k^t(\hat{z})$ for all $(\hat{z}, k, t)$ triple that υ_t maps in the set Φ_t.

        end for

    end for

    for all $\hat{z} \in \widehat{\mathcal{Z}}$ **do**

        Compute $u_{\hat{z}} = \sum_{k\in K} \pi_k \left(c_k^\top \hat{z} + \sum_{t\in\mathcal{T}} f_k^t(\hat{z})\right)$.

    end for

    if $UB > \min\limits_{\hat{z}\in\widehat{\mathcal{Z}}} u_{\hat{z}}$ **then**

        $UB = \min_{\hat{z}\in\widehat{\mathcal{Z}}} u_{\hat{z}}$, $z^* = \mathrm{argmin}_{\hat{z}\in\widehat{\mathcal{Z}}} u_{\hat{z}}$.

    end if

end while

Algorithm 3 can be implemented in a parallel fashion. In the lower bounding step, scenario subproblems can be solved in parallel as the problems are independent, and it suffices to collect the candidate solutions at the end. Similarly, for the upper bounding step, given a maintenance schedule each scenario subproblem is time decomposable with respect to the maintenance epochs. The decomposed subproblems for each time period with respect to the each generator status can be solved in a distributed framework.

Alternative cuts

Starting from the second iteration of Algorithm 3, the previously explored solutions are discarded from the set of feasible solutions in step 6 by utilizing the binary nature of the first stage variables. Given a solution $\hat{z}$, we can remove it from the solution set by adding 'no good' cut of the form, as discussed in [28],

$$\sum_{(i,t):\hat{z}_{i,t}=1} (1 - z_{i,t}) + \sum_{(i,t):\hat{z}_{i,t}=0} z_{i,t} \geq 1, \tag{2.9}$$

to the original set of constraints of the scenario subproblems. In our problem, every generator can enter maintenance only once during the planning horizon, which is ensured by the constraint (2.2d). We can exploit this structure to improve (2.9) to the following cut

$$\sum_{(i,t):\hat{z}_{i,t}=0} z_{i,t} \geq 1. \tag{2.10}$$

Proposition 2. *Inequality* (2.10) *is valid and dominates the standard 'no good' cut (2.9).*

Proof. Let $\bar{S}_1(\hat{z}) := \{z : (2.2\text{d}), (2.9)\}$ be the set of all vectors satisfying the maintenance constraint (2.2d) and (2.9) for the solution $\hat{z}$, and let $\bar{S}_2(\hat{z}) := \{z : (2.2\text{d}), (2.10)\}$ be the set with (2.2d), and (2.10). Let $S_1(\hat{z}) = \bar{S}_1(\hat{z}) \cap \{0,1\}^{|T|\times|G|}$ and $S_2(\hat{z}) = \bar{S}_2(\hat{z}) \cap \{0,1\}^{|T|\times|G|}$. We first show that the linear programming (LP) relaxation of $S_1(\hat{z})$, namely $\bar{S}_1(\hat{z})$, is weaker than the LP relaxation of $S_2(\hat{z})$, namely $\bar{S}_2(\hat{z})$, i.e., $\bar{S}_1(\hat{z}) \supseteq \bar{S}_2(\hat{z})$.

Since $z_{i,t} \leq 1$ is implied by $\sum_{t\in\mathcal{T}} z_{i,t} = 1$, we have $\sum_{(i,t):\hat{z}_{i,t}=1}(1 - z_{i,t}) \geq 0$. Thus, $\sum_{(i,t):\hat{z}_{i,t}=1}(1-z_{i,t}) + \sum_{(i,t):\hat{z}_{i,t}=0} z_{i,t} \geq \sum_{(i,t):\hat{z}_{i,t}=0} z_{i,t}$, which implies that $\bar{S}_1(\hat{z}) \supseteq \bar{S}_2(\hat{z})$.

Next, we show that $S_1(\hat{z}) = S_2(\hat{z})$. Since $S_1(\hat{z}) \supseteq S_2(\hat{z})$ follows from the previous claim, it suffices to show $S_1(\hat{z}) \subseteq S_2(\hat{z})$. For this purpose, we show that only $\hat{z}$ is removed from the solution space and not the other possible feasible solutions. Suppose there exists z such that $\sum_{(i,t):\hat{z}_{i,t}=0} z_{i,t} \leq 0$, hence $z_{i,t} = 0$ for all $\hat{z}_{i,t} = 0$. As $\sum_{t\in\mathcal{T}} z_{i,t} = 1 = \sum_{t:\hat{z}_{i,t}=1} z_{i,t} + \sum_{t:\hat{z}_{i,t}=0} z_{i,t}$ for all $i \in \mathcal{G}$, this implies that $z_{i,t} = 1$ for all $\hat{z}_{i,t} = 1$. Thus, $z = \hat{z}$. Combining the above, we proved that (2.10) is stronger than (2.9) for the given formulation. □

Eliminating redundant transmission line constraints

We propose an enhancement by identifying the lines that never violate the transmission line constraint in (2.2j) by solving auxiliary linear programs, similar to [29]. For this purpose, as a preprocessing step, we solve an appropriate linear programming relaxation of the operational problem by maximizing the amount of flow on each line l subject to the worst case nodal demands and operational decisions. We then check whether the resulting flow is larger than the corresponding line's flow capacity, which allows us to identify the provably redundant constraints to eliminate from the model. We note that different analyses for removing inactive line constraints are proposed in the literature, such as for composite system reliability evaluation in [30], and for security constrained unit commitment in [31].

2.4 Computational Results

In this section, we provide our computational results on the WSCC 9-bus instance [22], 39-bus New-England Power System [23], and 118-bus instance [24]. The algorithm is implemented in Python using Gurobi 6.5 as the solver with Intel Xeon E5-2670 machine. We study a one-year maintenance plan with monthly maintenance decisions. For operations scheduling, we have daily operations in the joint problem, and hourly decisions in the evaluation phase. We assume that $Y_p = 1$, and $Y_c = 2$ months and $L = 2$. For the chance constraint (2.2b), experiments are conducted by setting ρ as $\lfloor |\mathcal{G}|/3 \rfloor$ with $\epsilon = 0.05$ and 0.10. This means that with a probability of at least $(1 - \epsilon)$ at most one third of the generators can enter corrective maintenance due to a failure. The proposed safe approximation in Section 2.2.2 is used for representing the chance-constraint. To illustrate the size of the problem instances, number of variables and constraints in the optimization model (2.2) for 50 scenarios are reported in Table 2.1.

Table 2.1: Size of the instances.

Number of	9-bus	39-bus	118-bus
First-stage binary variables	36	120	228
Second-stage binary variables	162000	540000	1026000
Second-stage continuous variables	216600	882600	2466600
Constraints	729604	2964611	9165620

A database of historical degradation signals generated by a rotating machinery application is used to reproduce generator degradation. This setup is explained in detail in [20], and also used in [4] for representing the generators‘ accumulated decays. We utilize these signals to estimate the RLDs of the generators. To generate failure scenarios, the RLD of each generator is discretized in monthly periods ($d = 12$), and an i.i.d. sample of scenarios are generated as discussed in Algorithm 1. Our goal is to present the computational results and demonstrate the advantages of the proposed solution methodology along three dimensions:

1. We demonstrate the computational efficiency of the proposed algorithm by comparing the solution times with the generic scenario decomposition algorithm [21]. We illustrate the speedups with parallel implementation.

2. We provide SAA analysis for constructing confidence intervals on the true optimal value and determining sufficient sample sizes for the presented test cases.

3. To demonstrate the relevance of a stochastic modeling approach in maintenance schedul-

ing, we compare the proposed stochastic program with the failure scenarios and the chance-constraint (2.2), which we refer to as the chance-constrained model with failure scenarios (CCMFS) with two other simplified models: i) a deterministic model (DM), and ii) a chance-constrained model (CCM). The DM formulation ignores the unexpected failures by not considering scenarios or the chance constraint. CCM is a simplified stochastic formulation that only uses the chance constraint without incorporating the failure scenarios.

2.4.1 Computational efficiency

In this section, we demonstrate the computational performance of our approach in three fold. Firstly, we illustrate the individual impact of each proposed algorithmic enhancement. Secondly, we examine the performance of our approach compared to the generic scenario decomposition algorithm [21] under different number of scenarios. We conclude our computational analysis by illustrating the computational gains due to the distributed framework. We demonstrate these results on a sample 9-bus case when $\rho = 1$ and $\epsilon = 0.10$. All instances are solved to a relative optimality tolerance of 0.5%.

Effect of algorithmic enhancements

In Table 2.2 we demonstrate the computational gain of each proposed improvement over the generic algorithm [21] for an instance with 50 scenarios using a single processor. We observe an overall speedup of approximately 6 times using all enhancements. Since decomposing the operational problem into smaller subproblems and identifying generator statuses help in minimizing the number of resolves, the enhancement related to time decomposability and generator statuses presents the most gain.

We note that identifying generator statuses and prescreening transmission line constraints require additional computational efforts. The computational effort of these operations are considered within the run time of these algorithms. Although the run times of these additional operations are negligible, they contribute significantly to the improved performance of our approach.

Table 2.2: Computation gain of each algorithmic enhancement.

	Run time (sec)	Speed-up
Generic Algorithm [21]	513.62	
With stronger cut (2.10)	503.33	x1.02
With time decomposability and status	165.03	x3.11
With transmission line preprocess	255.41	x2.01
With all enhancements (Algorithm 3)	85.79	x5.99

Effect of number of scenarios

We demonstrate the computational efficiency of the proposed solution methodology by comparing its performance with the generic scenario decomposition algorithm [21]. Figure 2.2 presents the computational performance with samples of 50, 100, 150, 200 scenarios. Run time increases sublinearly with respect to the sample size for the proposed algorithm, which is at a much higher rate in the generic algorithm. As sample size increases, the proposed algorithm becomes even more advantageous. We note that the run times reported for both approaches are with a single processor for illustrating only the effect of the algorithmic enhancements.

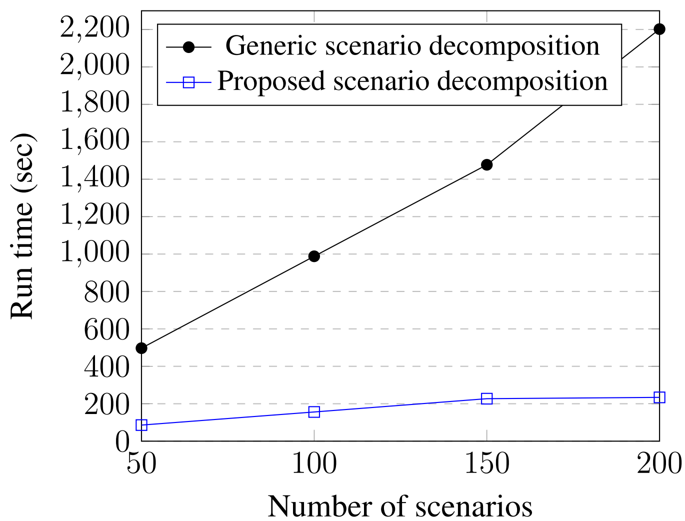


Figure 2.2: Effect of the proposed enhancements.

Effect of parallelism

In our experiments, we use a distributed framework in which the proposed scenario decomposition algorithm is parallelized. To evaluate the impact of parallelization, the speedup of the proposed algorithm is evaluated using different numbers of processors for solving the 9-bus instance with a sample of 500 scenarios (see Figure 2.3). The speed up ratio increases almost linearly until 8 processors. After 16 processors, the effect of parallelism begins to decrease. Nevertheless, the speedup is 134.83 times compared to the original algorithm when 32 processors are used. Note that the run time comparison is only presented for the 9-bus instance as the results are analogous for other test cases.

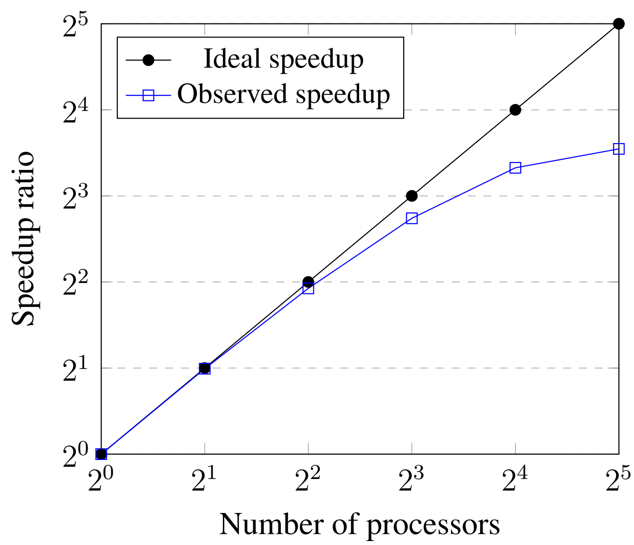


Figure 2.3: Effect of parallelism.

2.4.2 Sample average approximation results

In this section, we illustrate the two-step procedure to solve the stochastic program (2.2). Firstly, we apply SAA approach to construct statistical confidence intervals to the true objective due to the sampling of scenarios. Secondly, we apply the proposed scenario decomposition algorithm (Algorithm 3) to solve the resulting programs over the scenario subsets.

We present the confidence intervals for upper and lower bounds of the true optimal objective values to the proposed stochastic optimization models. To do so, we apply SAA method to the presented instances with $M = 5, N = 50, N' = 500$ when $\epsilon = 0.05, 0.10$. The resulting 95% confidence intervals for the lower bound and upper bound for 9-bus, 39-bus and 118-bus instances are presented in Table 2.4. The first column "rhs" gives the right hand side obtained by Proposition 1. In the last column of the table, we report the percentage gap between the lower and upper ends of the confidence interval to the true optimal in (2.8). As the true problem contains exponentially many scenarios in the order of number of generators, sampling becomes critical for ensuring computational tractability. Consequently, the gap values in Table 2.4 represent the statistical gaps estimated using the

SAA method described in Algorithm 2. Additionally, the objective function values of each of the 5 replications obtained using Algorithm 3 at the end of a two hour time limit are reported in Table 2.3.

Table 2.3: SAA objective function values for 5 replications (Results are in million \$).

Instance	ϵ	1	2	3	4	5
9-bus	0.05	2.48	2.48	2.45	2.45	2.45
	0.10	2.48	2.48	2.45	2.45	2.45
39-bus	0.05	82.02	81.93	81.99	81.95	81.88
	0.10	81.87	81.79	81.99	81.86	81.80
118-bus	0.05	57.91	57.34	57.30	57.21	57.43
	0.10	58.00	57.46	57.34	57.23	57.39

For both ϵ values, the resulting stochastic programs give the same solution in the 9-bus instance as the chance constraint restricts the feasible space considerably for small-scale instances. Our computational studies suggest that $N = 50$ provides sufficient accuracy for solving the stochastic program for the 9-bus instance, which has an extensive scenario size of $13^3 = 2197$. We also note that, σ_U is notably large as a failure of a single generator affects the objective value significantly in small-scale instances.

Table 2.4: SAA results (Intervals are in million \$).

	rhs	ϵ	CI for LB	CI for UB	Gap (%)
9-bus	0.05	0.05	(2.43 , 2.49)	(2.41 , 2.48)	2.15
	0.10	0.10			
39-bus	0.55	0.05	(81.88, 82.03)	(81.92, 82.12)	0.30
	0.72	0.10	(81.74, 81.98)	(81.83, 82.03)	0.36
118-bus	2.05	0.05	(57.00, 57.87)	(57.63, 57.99)	1.74
	2.41	0.10	(57.02, 57.95)	(57.63, 58.04)	1.79

Table 2.5: Model comparison for all instances (Estimates are in million \$).

		Evaluated Solution Estimates $(\hat{\mu}, \hat{\sigma})$			Cost Improvement (%)		p-value
		DM	CCM	CCMFS	DM-CCM	DM-CCMFS	CCM-CCMFS
9-bus	$\epsilon = 0.05$	(2.58, 0.87)	(2.41, 0.44)	(2.41, 0.44)	7.17	7.17	N/A
	$\epsilon = 0.10$		(2.50, 0.64)	(2.41, 0.44)	3.33	7.17	2.70e-06
39-bus	$\epsilon = 0.05$	(81.85, 2.11)	(75.61, 1.06)	(75.32, 1.04)	8.26	8.67	0
	$\epsilon = 0.10$		(75.44, 1.35)	(75.25, 1.03)	8.49	8.77	1.65e-04
118-bus	$\epsilon = 0.05$	(65.49, 3.74)	(55.09, 2.15)	(55.06, 1.91)	18.86	18.95	0.29
	$\epsilon = 0.10$		(55.13, 2.42)	(55.13, 2.17)	18.78	18.78	0.49

To study the performance in larger instances, we extend our results to 39-bus and 118-bus cases. Percentage gap indicates that $N = 50$ is sufficient for solving these instances to within 2% optimality for both ϵ values; instead of solving the extensive case with 13^{10} and 13^{19} many scenarios, respectively. As the decision maker becomes less conservative regarding the chance constraint, that is as ϵ value gets larger, a less expensive maintenance plan is obtained.

2.4.3 Model comparison

In this section, we highlight the importance of the stochastic formulation by comparing the quality of the maintenance schedules of the deterministic model, and the proposed chance-constrained models in Table 2.5. The column "Evaluated solution estimate" presents the performance of the schedules using 500 failure scenarios. The results are acquired by the following procedure: i) the optimal solutions to the DM, CCM and CCMFS for $N = 50$ with $\epsilon = 0.05, 0.10$ are obtained with respect to the aggregate operations problem, ii) the resulting maintenance schedules are evaluated using the detailed operational problem by fixing the first stage decision z and solving the second stage problem (2.6) for each scenario, and iii) the mean of the objective values of each scenario subproblem is calculated. CCMFS is solved over $M = 5$ batches and the solution with the best evaluated estimate is reported. The "Cost Improvement" column provides the percentage improvements gained by solving the stochastic programs. DM-CCM and DM-CCMFS columns give the percentage improvement gained by solving CCM instead of DM and CCMFS instead of DM, and reporting the percentage gap between the corresponding evaluated solution estimates, respectively. The "p-value" column represents the results of the paired t-tests for determining whether CCMFS outperforms CCM. In the paired t-test, the objectives of the 500 failure scenarios from the solutions of CCM and CCMFS are compared in a pairwise manner. The null hypothesis of the test is that CCM outperforms CCMFS, which corresponds to smaller objective values. The p-values are reported for a one-sided test. Note that p-value is not applicable for 9-bus instance with $\epsilon = 0.05$ since CCM and CCMFS arrive at the same solution.

The comparison of DM and CCMFS provides a Value of the Stochastic Solution (VSS) analysis. VSS compares the solutions of a deterministic model under a specific scenario, and a stochastic model. As we examine the effect of considering unexpected failure scenarios, we study the DM under the non-failure case in our analysis. Our analysis demonstrate 7-19% cost savings depending on the instance as can be seen in Table 2.5.

Since CCMFS captures more uncertainty using the chance constraint and the scenarios, the resulting solutions have less variance when evaluated under 500 scenarios, compared to the solutions obtained by DM and CCM. Thus, CCMFS results in more robust solutions that take into account various failure cases with less disruption to the maintenance plan. The improvement percentages in Table 2.5 indicate that the stochastic programs, CCM and CCMFS, are critical for all instances since the solutions found by the DM gives significantly higher objective values when it is evaluated under different failure scenarios. This demonstrates that unexpected failures should be considered explicitly when scheduling maintenance routines and determining operations. Finally, the p-values over 9-bus and 39-bus instances indicate that we have significant evidence to reject the null hypothesis. Thus, the proposed CCMFS approach performs better than CCM. For 118-bus instance, we cannot reject the null hypothesis, thus CCM can be preferred for this case due to its computational efficiency.

2.5 Concluding Remarks

This chapter presents a novel framework for solving the joint maintenance and operations scheduling problem by considering generator failures. We leverage on degradation-based predicted RLDs to compute maintenance costs and generator failure probabilities, which are then integrated into a stochastic mixed-integer optimization model that determines optimal maintenance and operational decisions. We present a chance-constraint that adapts to the generator RLDs in order to restrict the number of generators that enter maintenance due to a failure with high probability. We derive a deterministic safe approximation for this chance constraint. We develop a scenario decomposition algorithm by introducing various enhancements and combine it with a sampling approach to solve the stochastic optimization model. Our experiments show significant computational gains over generic scenario decomposition and serial implementations. Finally, we demonstrate that the proposed approach provides significant improvements over the models that use proxy cost functions, demonstrating the importance of considering unexpected generator failures while scheduling maintenance and operations.

CHAPTER 3

DATA-DRIVEN MAINTENANCE AND OPERATIONS SCHEDULING IN POWER SYSTEMS UNDER DECISION-DEPENDENT UNCERTAINTY

3.1 Introduction

Ensuring reliable and cost effective operations of generators is an important problem in power systems. One of the key factors that impacts this problem is related to generator maintenance scheduling, since resulting schedules determine availability and reliability of the generators. Most classical maintenance approaches rely on predetermined time intervals or safety margins for scheduling maintenance. Consequently, these methods often result in excessive or unnecessary maintenance events, especially when designed in a conservative manner. They also do not provide much needed visibility into the actual condition of the generators, and thus, still experience unexpected failures. The emergence of sensor technology has enabled the incorporation of generator state-of-health into maintenance and operations scheduling. Operational decisions, such as on-off actions for generators and their dispatch amount, play a pivotal role in this regard as higher (lower) loads result in an accelerated (decelerated) degradation process, which may require scheduling maintenance at an earlier (or later) time. Thus, the effect of operational decisions on the corresponding component's aging is critical in determining its availability. To address these issues, we propose an optimization framework for condition-based generator maintenance and operations scheduling that accounts for the loading profiles derived from the operational decisions.

Generator maintenance scheduling constitutes an important class of problems in power systems, see [7] for a recent survey. The literature on maintenance scheduling can be categorized into three groups. The first group of studies focuses on periodic maintenance routines that are based on a predetermined time schedule ([8, 12]) with additional constraints that inform maintenance decisions, such as enforcing one maintenance per year. These studies neglect potential information related to the condition/performance of the generators. The second line of research [9, 32, 33] adopts a reliability-based approach by considering failure rates and reliability metrics, such as mean-time-to-failure for scheduling maintenance. These approaches tend to adopt a general schedule across all generators regardless of their unit-to-unit variations, whether in the way they are operated or the manner in which they degrade. A third group of recent studies [4, 5, 1] considers generator degradation by monitoring cumulative damage and other forms of wear and tear using sensor technology. These studies focus on leveraging real-time generator-specific degradation signals to estimate statistical distributions of the generator's remaining lifetime. The predictions can be updated in real-time and leveraged when solving the operational problem. However, these models stop short of modeling the impact of operational decisions on generator degradation (hereafter referred to as *load-independent* models). In reality, generator loading has a significant effect on how fast a generator degrades. Harsh operating and loading conditions, i.e., dispatch decisions, tend to accelerate physical degradation. Thus, there is a tight coupling between the operational decisions and its effect on generator degradation.

Data-driven degradation modeling has played a key role in predicting remaining lifetime of machines and capital-intensive assets. A review of various modeling approaches used for this purpose can be found in [34]. The relationship between loading conditions and equipment reliability has been a well-studied area in the reliability literature. Traditionally, operating conditions have been considered as model covariates in classic location-scale models and proportional hazard models, see [35]. The accelerated degradation testing framework proposed by [36] is among the early models that considers the interaction between the stress conditions and degradation rates. The authors model degradation using a Brownian motion and present a time-scale transformation proportional to the testing stress. Recently, time-varying operating conditions are integrated with sensor-based degradation models for predicting remaining life distributions in an adaptive manner using a Bayesian framework, see [37, 38, 39].

Although load-dependent degradation modeling has been studied in the reliability literature, integrating it into maintenance optimization has not been explored in depth. Ideally, effective maintenance scheduling should not only consider the health of an asset, but also its loading/operating conditions. In power systems, operating conditions of the generators are determined by the unit commitment (UC) problem, which is a well-studied problem in literature (see [16, 40] for recent reviews). The UC problem can be also referred as the operational problem in power systems. This problem aims to determine which generators will be on or off and how much energy they need to produce by considering restrictions of the underlying power network such as generation capacities, transmission line limits, etc. As the loading/operating conditions are critical in assessing generators' conditions, the operational problem and corresponding load-dependent degradation modeling need to be incorporated into maintenance scheduling. With the exception of recent work by [41], there has not been a formal framework for using load-dependent degradation models in scheduling maintenance. In [41], the authors solve the UC problem by considering the effects of generator loads on degradation. They categorize generator loading into three levels and jointly optimize operations and maintenance decisions. Although the authors consider the stochasticity of the degradation process, they neither formally account for unexpected failure possibilities nor the continuous load amounts within the optimization model.

Most of the power system problems involve uncertainties which can only be addressed by a stochastic optimization framework. These uncertainties also play a critical role in the joint scheduling of maintenance and operations. Several studies [17, 18, 19] consider uncertainty of price and demand for this scheduling problem. Recall that Chapter 2 proposes a stochastic optimization framework for modeling sensor-based generator maintenance and operations scheduling problem by explicitly considering generator failures through scenarios and chance constraints. Nevertheless, the proposed approach does not consider the effect of the operational decisions on generators' degradation. Neglecting this effect results in a simplified prognostics procedure and a mixed-integer linear optimization model to determine maintenance and operations schedules. Modeling the dependency between operational decisions and degradation processes involve decision-dependent (endogenous) uncertainties. In decision-dependent uncertainty, the distribution of a random variable changes due to the planner's decisions, and complex modeling procedures are required for its correct representation (see [42]). In our case, the dispatch levels of generators affect the distribution of the generators' remaining lifetimes, resulting in an endogenous uncertainty to represent unexpected failure possibilities, which need to be considered when scheduling maintenance and operations.

In this chapter, we develop a novel stochastic optimization framework that determines the maintenance and operations schedules of the generators while considering the impact of the decision-based degradation amounts. We refer to our comprehensive approach as *load-dependent* in the rest of this chapter. Our main contributions can be summarized as follows:

1. We formulate a stochastic optimization model that jointly optimizes maintenance and operations while accounting for the endogenous effect of operational decisions on degradation. The resulting model includes nonlinearities due to the decision-dependent structure of the cumulative distribution functions of the remaining lifetimes and maintenance cost functions associated with generators. We develop linearization procedures for these nonlinearities to obtain a stochastic mixed-integer linear programming formulation, and propose formulation enhancements. We also extend the chance constraint proposed in Chapter 1 to the decision-dependent setting.

2. We build on existing literature and propose a data-driven degradation modeling framework for capturing the effect of the continuous loading profiles (rather than discrete levels) that are functions of the generators' minimum and maximum production capacities. We also develop an estimation method to consider the effect of the load decisions by taking into account the signal variability.

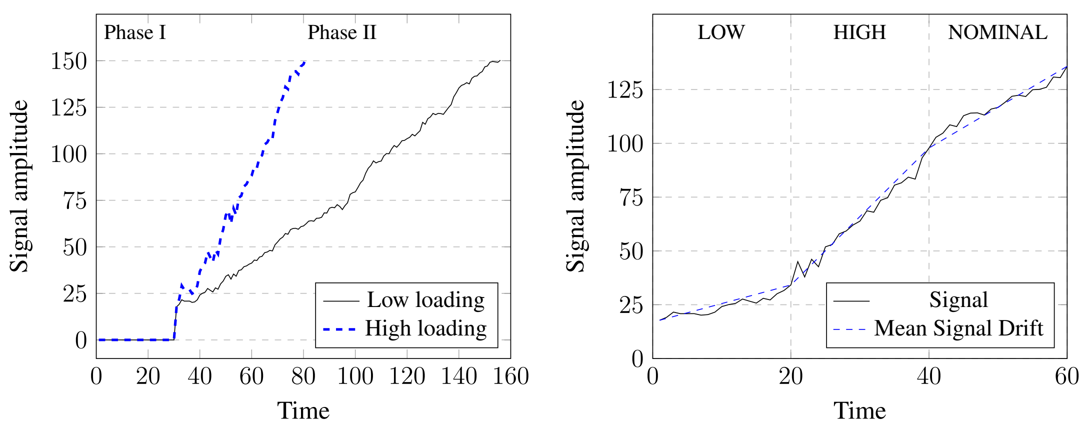
3. We propose a decision-dependent simulation framework to evaluate the performances of resulting maintenance and operations schedules. We provide a comprehensive computational study on representative IEEE test cases by comparing the proposed load-dependent approach with load-independent approaches under different congestion levels and conservativeness amount of the chance constraint. Our experiments demonstrate the success of load-dependent schedules resulting in significant cost savings of up to 20% and failure preventions. Finally, we provide experiments demonstrating the computational efficiency of the formulation improvements compared to generic methods with more than 17 times speedup.

The remainder of this chapter is organized as follows: Section 3.2 discusses the degradation modeling framework along with the data-driven estimation procedures. Section 3.3 presents the optimization methodology for modeling load-dependent condition-based maintenance and operations schedules with our proposed enhancements. Section 3.4 illustrates the computational results by developing a simulation procedure, and the efficiency of the proposed approach from various aspects. Section 3.5 concludes the chapter with final remarks.

3.2 Degradation Modeling

Unexpected failures of power generators can have catastrophic consequences. Thus, monitoring the state-of-health of power generators is necessary to maintain their availability and improve their reliability. Condition monitoring, the process of using sensors to assess the state-of-health of a machine, is becoming more prevalent across numerous industrial applications. Raw sensor data from power generators can be synthesized into degradation signals that represent the severity of the underlying physical degradation processes taking place in the generator. The degradation state of the generators and their remaining lifetime distributions can be estimated using degradation signals.

One of the most critical factors that determines the degradation of a generator is severity of its operating condition. Generators tend to degrade faster when operated close to their maximum capacity. To demonstrate this effect, we illustrate degradation signals in Figure 3.1 that follow a Brownian motion with drift to mimic the degradation of a rotating machinery under varying loading conditions. Figure 3.1a shows an example of two degradation signals under low and high loading. Phase I is considered an "as good as new" state with no signs of degradation whereas Phase II highlights the progressive nature of physical degradation and its manifestation in a gradually increasing degradation signal. Figure 3.1b represents potentially how the degradation rate changes with low and high loading conditions. The dotted line in Figure 3.1b shows the mean drift of the degradation signal under each load condition. The plot illustrates how the degradation rate at low loading from time 0 to 20 is lower than the nominal condition. The converse is true for the high load from time 20 to 40. This example highlights the importance of capturing the effects of operational decisions on generator conditions when solving the maintenance and operations scheduling problem.



(a) Effect of load decisions on two signals.

(b) Effect of time varying load levels.

Figure 3.1: Illustration of the effect of load decisions on degradation signals.

Usually, an equipment operates for a period of time without any signs of degradation (Phase I). This phase is often random and hypothetically equipment does not fail in this period due to degradation. However, equipment is often subject to external factors such as human related operational errors that can cause an unexpected failure event. These failures are rare and not related to equipment's degradation, and often cannot be predicted. Phase II is characterized by a degradation process that is significant enough to be observed using some form of sensor technology. Most machines and equipment can operate in these partially degraded modes for a significant period of time. In fact, Phase II is typically where degradation trends can be leveraged when predicting remaining lifetime. Our analysis focuses on the second phase of a degradation process for modeling failures due to degradation.

3.2.1 Load-dependent degradation model

We model the degradation signal of a generator i as a continuous-time continuous-state model, denoted by $\{S_i(t) : t \geq 0\}$. We assume that $S_i(t)$ has the following functional form;

$$S_i(t) = \theta_i + \nu_i t + \sigma_i W(t), \tag{3.1}$$

where θ_i is the initial signal amplitude and ν_i is the drift of the process (see Figure 3.1b for a sample signal). The value of σ_i corresponds to standard deviation of the signal, which is assumed to be known and same across the population of generators, denoted by σ for all generators. The process $W(t)$ represents a standard Brownian motion with linear drift, where $W(0) = 0$. The increments $W(t+u) - W(t)$ for $u \geq 0$ are independent and identically distributed and follow a Normal distribution with mean 0 and variance u. We adopt the form (3.1) for degradation modeling as it is widely used in real-time condition monitoring (see [20]).

To model different degradation rates that correspond to the loading levels, we use the time transformation approach proposed by [36] where the notion of *effective time* is used to scale the time under each stress condition. We extend the notion of effective time as follows.

Definition 1. *The effective time of generator i at time t is defined as $\tau_i(t) := \int_0^t L_i(t')dt'$, where the function $L_i(t)$ represents the load level of generator i.*

Here, the value of the load function $L_i(t)$ can be interpreted as a *load multiplier*. The value of this function is equal to 1 under nominal loading, which increases or decreases based on the dispatch decisions of the corresponding generator. In other words, the load function $L_i(t)$ is used to scale the signal $S_i(t)$ based on the loading condition. The scaled time is denoted by $\tau_i(t)$. In the case of nominal loading, $\tau_i(t) = t$. By using $\tau_i(t)$ in Equation 3.1, we can rescale the time period of different segments of the degradation signal by their respective load multiplier, thus allowing us to recreate a corresponding degradation signal with constant drift. It is noteworthy to mention that for the load-independent case $\tau_i(t)$ equals to t irrespective of the dispatch (loading) decisions, which may lead to inaccurate remaining life predictions.

3.2.2 Load-based remaining life estimation

Our underlying assumption is that different loading regimes will have the same effective time if their cumulative degradation is equivalent. In particular, we assume that a failure occurs when the degradation level of generator i, $S_i(t)$, reaches a predefined threshold value, Λ. In general settings, the parameters of the signal model (3.1) are unknown, and need to be estimated.

We assume that the unknown model parameters follow a prior distribution that can be estimated from historical data. The prior distribution represents the characteristics of the generator's population. The key assumption here is that the degradation process of the population exhibits a common functional trend. Specifically, we denote the prior distributions of θ_i and ν_i as $\pi_1(\theta_i)$ and $\pi_2(\nu_i)$. The prior distributions are assumed to follow a Normal distribution with mean μ_0 and variance σ_0^2, and mean μ_1 and variance σ_1^2, respectively. The random variables θ_i and ν_i are assumed to be mutually independent. The prior distribution will be updated using real-time signals observed from each generator using a Bayesian framework similar to the one proposed in [20]. This allows the model to adapt to the unique degradation characteristics of each generator resulting in remaining life predictions that are driven by the generator's degradation process. To see this, assume that the signal levels at times $t_1^i, t_2^i, \cdots, t_k^i$ for every generator i are observed as $S_i(t_1^i), \ldots, S_i(t_k^i)$. As highlighted by Equation 3.1, we assume that the degradation signal follows a Brownian motion. Thus, we focus on modeling the increments of the signal, which we denote as $S_j^i = S_i(t_j^i) - S_i(t_{j-1}^i)$ for $j = 2, \cdots, k^i$, where $S_1^i = S_i(t_1^i)$. We assume that the future loading function of each generator i, $\{L_i(t) : t \geq 0\}$, is known a priori. By adopting a Bayesian updating approach (Proposition 2 in [20]) and utilizing the effective time notion (Definition 1), we can find the posterior distributions of θ_i and ν_i.

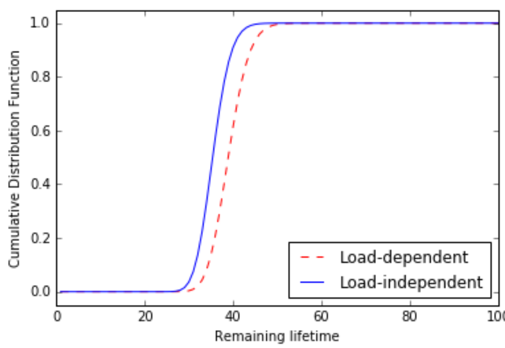
Proposition 3. *Given the observed signal increments, S_j^i, $j = 1, \cdots, k^i$, with parameters (ν_i, θ_i), and failure threshold Λ; for a load function, $\{L_i(t) : t \geq 0\}$, the posterior mean of the drift parameters ν_i, is given by,*

$$\mu_i' = \frac{(\sigma_1^2 \sum_{j=1}^{k^i} S_j^i + \mu_1 \sigma^2)(\sigma_0^2 + \sigma^2 t_1'^i) - \sigma_1^2 (S_1^i \sigma_0^2 + \mu_0 \sigma^2 t_1'^i)}{(\sigma_0^2 + \sigma^2 t_1'^i)(\sigma_1^2 t_k'^i + \sigma^2) - \sigma_0^2 \sigma_1^2 t_1'^i},$$

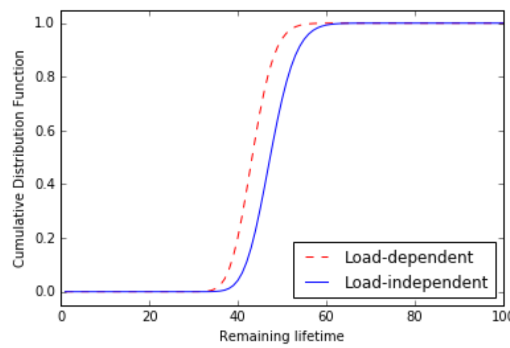
where $t_j'^i = \int_0^{t_j^i} L_i(t') dt'$, for $j = 1, \cdots, k$. Then, the corresponding remaining lifetime at t_k^i has an inverse Gaussian distribution $IG(w|\phi, \upsilon)$, where $w = \tau_i(t) = \int_0^t L_i(t_k^i + t') dt'$, $\phi = \frac{\Lambda - \sum_{j=1}^{k^i} S_j^i}{\mu_i'}$, and $\upsilon = \frac{(\Lambda - \sum_{j=1}^{k^i} S_j^i)^2}{\sigma_i^2}$.

Combining the Bayesian update procedure with the notion of effective time, we can compute updated load-dependent remaining lifetimes. We note that for the load-independent models, Proposition 3 can also be used for estimating the remaining lifetime distribution of signal i by replacing $t_j'^i$, with actual time t_j^i for $j = 1, \cdots, k$, and using $t_k^i + t$ in place of the effective time, $\tau_i(t)$.

To illustrate the difference in remaining life estimation between the load-dependent and load-independent approaches, we examine two specific cases. Consider a case where generator i is consistently operated under a high loading level until the k^{th} observation epoch, i.e., operating time, t_k^i. The effective time of this generator will be greater than the observed time, i.e., $t_k'^i > t_k^i$. Next, assume that the loading condition are switched and the generator will operate under a nominal load level for the rest of its lifetime. In a load-independent case, the remaining life distribution will be underestimated since it is based on an *inflated* drift parameter that assumes that the prevailing loading conditions remain the same. In contrast, the load-dependent model utilizes a drift value that has been adjusted based on the future loading level, nominal load. Figure 3.2a highlights the difference in the estimated cumulative distribution function of the remaining lifetime on a set of simulated signals using the two kinds of modeling approaches. The converse is also true. Figure 3.2b highlights the case where a generator operates at a less than nominal loading condition, i.e., $t_k'^i < t_k^i$, which overestimates the remaining lifetime once the future load function increases to nominal load.



(a) The case when $t_k'^i > t_k^i$.



(b) The case when $t_k'^i < t_k^i$.

Figure 3.2: CDFs of remaining lifetime for load-dependent and load-independent approaches.

3.2.3 Effective time estimation

In order to reflect the effects of operational decisions on degradation, one needs to accurately map the relationship between the load function and the decisions evaluated by the optimization model. In particular, the value of the load multiplier function $L_i(t)$ depends on the dispatch and maintenance decisions of generator i during period t while taking into account minimum and maximum production capacities. In this section, we discuss how we model the effective time based on the operational decisions and degradation uncertainty.

As mentioned before, we assume that load levels in each period are known, however drift values of the signals are uncertain. Our framework accommodates continuously varying load functions. Power generators, however, often operate under a fixed load level for a pre-specified period of time depending on their operating schedule. Consequently, it is reasonable to assume that load levels remain the same in between consecutive signal observations as the time of the signal observations correspond to the operational periods.

To estimate the effective time, we consider a set of historical degradation signals, namely $\mathcal{I}$, associated with generators that have been observed until their time of failure. In this context, our focus is on *soft failures* defined by unacceptable or alarming generator performance as opposed to *hard failures* that often result from catastrophic damage. We observe each signal $i \in \mathcal{I}$ until its failure time T_i, and consider the load function at the discrete time points from $0, 1, \cdots, T_i$. For $|J|$ levels of load, namely $L_1, L_2, \cdots, L_J$ and signal $i \in \mathcal{I}$, the mean estimate of the signal drift parameter, denoted by $\overline{\mu}_{i,j}$, can be estimated using the following expression:

$$\overline{\mu}_{i,j} = \frac{\sum_{t=1:L_i(t)=L_j}^{T_i} S_t^i}{\sum_{t=1}^{T_i} \mathbb{1}_{\{L_i(t)=L_j\}}}, \tag{3.2}$$

for $\sum_{t=1}^{T_i} \mathbb{1}_{\{L_i(t)=L_j\}} > 0$. In order to model the effective time, we estimate its value over each unit time. For this purpose, we normalize the mean estimates corresponding to each signal $i \in \mathcal{I}$ and load $j \in J$ pair, by dividing them the overall average signal drift, denoted by $\overline{\overline{\mu}}$. We denote the corresponding normalized estimates as $\overline{\mu}'_{i,j} = \overline{\mu}_{i,j}/\overline{\overline{\mu}}$.

We define the set of loading levels corresponding to the generator i as $L_{i,j} = p_i^{min} + (p_i^{max} - p_i^{min})(j-1)/(J-1)$ for $j = 2, \cdots, J-1$, and $L_{i,1} = p_i^{min}$, $L_{i,J} = p_i^{max}$, where p_i^{min} and p_i^{max} represent minimum and maximum production requirements of the corressponding generator. Using the $(L_{i,j}, \overline{\mu}'_{i,j})$ points, we develop a linear regression model to estimate the effective unit time, which we denote by $d_{i,t}$. We assume that generators have different capacities, and thus, we estimate an individual regression model for each generator. The resulting model for generator i can be represented as $d_{i,t} = \alpha'_i L'_i(t) + \beta'_i$, where $d_{i,t}$ is the estimated effective unit time at time t as a function of load level $L'_i(t)$, and α'_i and β'_i are the regression coefficients. We note that the load function $L'_i(t)$ is in terms of generation capacities. The relationship between the effective time $\tau_i(t)$ and $d_{i,t'}$ can be expressed as $\tau_i(t) = \sum_{t'=1}^{t} d_{i,t'}$.

To improve the practical relevance of our model, we assume that a generator does not degrade when it is not operating. Therefore, we integrate the commitment variable, $x_{i,t}$, into our regression model as follows; $d_{i,t} = \alpha'_i L'_i(t) + \beta'_i x_i$ where x_i is 1 when the generator operates and 0 otherwise. The regression models are foundational to characterizing the relationship between operational decisions and efficient maintenance scheduling of the generators.

3.3 Optimizing Maintenance and Operations

In this section, we formulate the load-dependent generator maintenance and operations scheduling problem as a decision-dependent stochastic program with cost and reliability perspectives. Given a fleet of generators, our aim is to obtain their maintenance and operations schedules while simultaneously minimizing maintenance and operations costs, and satisfying the system constraints under the load-dependency of generators' conditions. We consider a one-year planning horizon with monthly maintenance decisions, and daily operational schedules corresponding to commitment decisions, dispatch and demand curtailment amounts. We allow one maintenance per each generator during the planning horizon. Additionally, we consider a capacity limit on the number of ongoing maintenances. We note that a generator needs to be off if it is under maintenance. We also take into account operational level restrictions such as demand satisfaction, production capacities, and transmission line limits on the underlying power network.

Maintenance routines can be categorized into two groups. A preventive maintenance is conducted at the scheduled maintenance period, which costs C^p. Otherwise, a corrective maintenance is performed if a generator fails unexpectedly before its scheduled maintenance period with a cost, C^c. Corrective maintenance typically costs more and lasts longer compared to a scheduled maintenance. Thus, our aim is to identify cost effective and reliable maintenance and operations schedules that result in fewer number of unexpected failures with lower overall costs. To represent the trade-off between preventive and corrective maintenance, we adopt the *dynamic maintenance cost function* approach presented in [43, 5, 1]. The cost function uses the preventive and corrective maintenance costs coupled with the remaining life distribution of the generator to calculate the overall maintenance cost at future time epochs. We note that our framework enables updating the remaining lifetime estimations of the generators through newly acquired real-time degradation signals. This impacts the cost function, which is also dynamically revised to account for real-time changes in the degradation state of the generator. We extend the dynamic maintenance function definition to the load-dependent setting by integrating the effective time approach introduced in Section 3.2.1. We first define the decision variable $\tau_{i,t}$ as the effective age of generator i after t periods from the beginning of planning horizon of the optimization model. Next, the dynamic maintenance cost of generator i at time t with initial effective age $\tau_{i,0}$ can be expressed as follows:

$$C_{i,\tau_{i,0}}(\tau_{i,t}) = \frac{C^p \Pr(R_{i,\tau_{i,0}} > \tau_{i,t}) + C^c \Pr(R_{i,\tau_{i,0}} \leq \tau_{i,t})}{\int_0^{\tau_{i,t}} \Pr(R_{i,\tau_{i,0}} > z)dz \ + \tau_{i,0}}, \tag{3.3}$$

where $R_{i,\tau_{i,0}}$ is the remaining lifetime of generator i given the initial effective age $\tau_{i,0}$. We assume that the value of $\tau_{i,0}$ is known for every generator i at the beginning of planning.

Below is a summary of the sets, decision variables and parameters of the optimization model.

Sets:

$\mathcal{B}$ Set of buses.

$\mathcal{G}$ Set of generators.

$\mathcal{L}$ Set of transmission lines.

$\mathcal{S}$ Set of operational subperiods within a maintenance period.

$\mathcal{T}$ Set of maintenance periods in the planning horizon.

Decision variables:

$z_{i,t}$ 1 if generator i enters maintenance in maintenance period t, and 0 otherwise.

γ_t Additional maintenance capacity added in maintenance period t.

$\tau_{i,t}$ Effective age of generator i at time t.

$x_{i,t,s}$ 1 if generator i is on in operational period s of maintenance period t, and 0 otherwise.

$y_{i,t,s}$ Dispatch amount of generator i in operational period s of maintenance period t.

$\psi_{b,t,s}$ Demand curtailed at bus b in operational period s of maintenance period t.

Parameters:

C_{add} Per unit cost of maintenance capacity added.

$V_{i,t,s}$ No-load cost of generator i in the operational period s of maintenance period t.

$F_{i,t,s}$ Per unit dispatch cost of generator i in operational period s of maintenance period t.

P_{DC} Per unit cost of demand curtailed.

ξ Maintenance criticality coefficient.

H Planning horizon length in maintenance periods.

$\overline{M}$ Maximum number of ongoing maintenances.

Y_p Duration of a preventive maintenance.

ϵ Confidence level of the chance constraint.

ρ Threshold on the number of generators to fail.

$D_{b,t,s}$ Demand of bus b in operational period s of maintenance period t.

p_i^{min} Minimum production requirement of generator i.

p_i^{max} Maximum production capacity of generator i.

f_{max}^{l} Flow capacity of line l.

a_l Shift factor vector for line l.

$M_{b,i}$ 1 if generator i is on bus b, and 0 otherwise.

The resulting load-dependent generator maintenance and operations scheduling problem can be formulated in (3.4) as follows:

$$
\begin{aligned}
\min\ & \xi \left(\sum_{i \in \mathcal{G}} \sum_{t \in \mathcal{T}} C_{i,\tau_{i,0}}(\tau_{i,t}) z_{i,t} + \sum_{t \in \mathcal{T}} C_{add} \gamma_t \right) \\
& + \sum_{i \in \mathcal{G}} \sum_{t \in \mathcal{T}} \sum_{s \in \mathcal{S}} (V_{i,t,s} x_{i,t,s} + F_{i,t,s} y_{i,t,s}) + \sum_{t \in \mathcal{T}} \sum_{s \in \mathcal{S}} \sum_{b \in \mathcal{B}} P_{DC} \psi_{b,t,s} && \text{(3.4a)} \\
\text{s.t. }\ & \tau_{i,t} = \sum_{t'=1}^{t} \sum_{s \in \mathcal{S}} (\alpha_i y_{i,t',s} + \beta_i x_{i,t',s}) \qquad i \in \mathcal{G}, t \in \mathcal{T} && \text{(3.4b)} \\
& \text{Pr} \left(\sum_{i \in \mathcal{G}} \sum_{t \in \mathcal{T}} \zeta_{i,t}(\tau_{i,t}) z_{i,t} \leq \rho \right) \geq 1 - \epsilon && \text{(3.4c)} \\
& \sum_{e=0}^{Y_p - 1} z_{i,t-e} \leq \overline{M} + \gamma_t \qquad t \in \mathcal{T} && \text{(3.4d)} \\
& \sum_{t \in \mathcal{T}} z_{i,t} = 1 \qquad i \in \mathcal{G} && \text{(3.4e)} \\
& x_{i,t,s} \leq 1 - \sum_{e=0}^{Y_p - 1} z_{i,t-e} \qquad i \in \mathcal{G}, t \in \mathcal{T}, s \in \mathcal{S} && \text{(3.4f)} \\
& \sum_{i \in \mathcal{G}} y_{i,t,s} + \sum_{b \in \mathcal{B}} \psi_{b,t,s} = \sum_{b \in \mathcal{B}} D_{b,t,s} \qquad t \in \mathcal{T}, s \in \mathcal{S} && \text{(3.4g)} \\
& p_i^{min} x_{i,t,s} \leq y_{i,t,s} \leq p_i^{max} x_{i,t,s} \qquad i \in \mathcal{G}, t \in \mathcal{T}, s \in \mathcal{S} && \text{(3.4h)} \\
& \left| \sum_{b \in \mathcal{B}} a_{l,b} \left(\sum_{i \in \mathcal{G}} M_{b,i} y_{i,t,s} + \psi_{b,t,s} - D_{b,t,s} \right) \right| \leq f_l^{max} \qquad t \in \mathcal{T}, s \in \mathcal{S}, l \in \mathcal{L} && \text{(3.4i)} \\
& z_{i,t}, x_{i,t,s} \in \{0,1\}, \gamma_t, y_{i,t,s} \geq 0, D_{b,t,s} \geq \psi_{b,t,s} \geq 0 \quad i \in \mathcal{G}, t \in \mathcal{T}, s \in \mathcal{S}, b \in \mathcal{B}. && \text{(3.4j)}
\end{aligned}
$$

Objective function (3.4a) minimizes total maintenance and operations cost of a fleet of generators. The first part of the objective represents the maintenance cost, in particular the dynamic maintenance cost (3.3) and additional labor costs. To approximate the maintenance cost function, we propose a piecewise linearization approach, which is described in detail in Section 3.3.2. The remaining part of the objective corresponds to the operational costs including the costs of commitment, dispatch, and demand curtailment. The cost of maintenance is adjusted with respect to the cost of operations by the maintenance criticality coefficient, ξ. Selection of ξ values enables examining the importance of the maintenance and operations costs on the resulting schedules.

Constraint (3.4b) represents the effective age formulation. The modeling of this constraint and the derivation of the corresponding parameters (α', β') are described in detail in Section 3.2.3. In order to ensure that the variable $\tau_{i,t}$ represents the effective age of the generator i in terms of maintenance periods, the parameters (α_i, β_i) for each generator i are taken as $\alpha_i = \alpha'_i/|\mathcal{S}|$, and $\beta_i = \beta'_i/|\mathcal{S}|$. This change of parameters helps in establishing the time transformation from operational periods (i.e. days) to maintenance periods (i.e. months).

The chance constraint (3.4c) aims to restrict the number of generators that fail before their scheduled maintenance with a threshold ρ with high probability $1 - \epsilon$. This constraint leverages sensor information through the random variable ζ. The Bernoulli random variable $\zeta_{i,t}$ is 1 if $\tau_{i,t} \geq R_{i,\tau_{i,0}}$ and 0 otherwise. This constraint formulates a decision-dependent uncertainty, as the failure probabilities depend on $\tau_{i,t}$, which is related with the dispatch and commitment decisions through constraint (3.4b). As the chance constraint is computationally intractable, we develop a combination of safe approximation and piecewise linearization approaches for its representation in Section 3.3.1 and Section 3.3.2, respectively.

Constraint (3.4d) guarentees that there is at most $\overline{M} + \gamma_t$ maintenances in each period t. Thus, the maintenance capacity of the system can be violated in return for its penalty in the objective. Constraint (3.4e) ensures that each generator enters maintenance once through the planning horizon, which is a common assumption in generator maintenance scheduling in power systems literature (see e.g. [12], [6]). Constraint (3.4f) enforces the generators to be off if they are under maintenance.

The remaining constraints in (3.4) represent the operational level restrictions. In particular, constraint (3.4g) ensures that total demand is satisfied with production and demand curtailment. Constraint (3.4h) guarentees that generators produce within their production limits, and constraint (3.4i) enforces the transmission line limits by considering the DC approximation (see [44]) for modeling the power flow.

3.3.1 Safe approximation of the chance constraint

The proposed chance constraint (3.4c) poses computational challenges, as it is intractable to represent and considers decision-dependent uncertainty. For this purpose, we present alternative ways for reexpressing this constraint. We utilize a deterministic safe approximation of the chance constraint as follows:

Proposition 4. *The deterministic constraint*

$$\sum_{i\in\mathcal{G}}\sum_{t\in\mathcal{T}}\mathbf{E}[\zeta_{i,t}(\tau_{i,t})]\, z_{i,t} \leq \max\left(\rho\,\epsilon, \max_{\delta>0}\left[\frac{((\epsilon\, e^{\delta\rho})^{1/|\mathcal{G}|}-1)|\mathcal{G}|}{e^{\delta}-1}\right]\right) = \rho^{*}, \tag{3.5}$$

is a safe approximation of (3.4c)*, i.e. any* $z \in \{0,1\}^{|\mathcal{T}|\times|\mathcal{G}|}$ *satisfying* (3.5)*, satisfies* (3.4c).

We note that Proposition 4 is an extension of Proposition 1 in Chapter 1, in which the random variable $\zeta_{i,t}$ is independent of the effective age of the generator $\tau_{i,t}$.

The term $\mathbf{E}[\zeta_{i,t}(\tau_{i,t})]$ in Proposition 4 can be expressed as $\mathbf{E}[\zeta_{i,t}(\tau_{i,t})] = \Pr(R_{i,\tau_{i,0}} \leq \tau_{i,t})$, using the definition of the Bernoulli random variable ζ. To represent this decision-dependent uncertainty, we define an auxiliary decision variable $P_{i,t} := \mathbf{E}[\zeta_{i,t}(\tau_{i,t})]$. Considering $\bar{P}_{i,t}$ as an upper bound on $P_{i,t}$, and utilizing $0 \leq P_{i,t} \leq \bar{P}_{i,t} \leq 1$, we can linearize the term $v_{i,t} := P_{i,t}\, z_{i,t}$. The safe approximation of the chance constraint (3.4c) is represented in the form in (3.6).

$$\sum_{i\in\mathcal{G}}\sum_{t\in\mathcal{T}} v_{i,t} \le \rho^*, \tag{3.6a}$$

$$0 \le v_{i,t} \le P_{i,t},\;\; P_{i,t} - (1-z_{i,t})\bar{P}_{i,t} \le v_{i,t} \le \bar{P}_{i,t} z_{i,t}. \tag{3.6b}$$

Similarly, the objective function (3.4a) includes nonlinear terms. To handle this issue, we linearize $C_{i,\tau_{i,0}}(\tau_{i,t})z_{i,t}$. Let $\theta_{i,t}$ be $C_{i,\tau_{i,0}}(\tau_{i,t})$. Then, we define $w_{i,t} := \theta_{i,t} z_{i,t}$. Since the cost of corrective maintenance is an upper bound on the dynamic maintenance cost function, we observe that $0 \le \theta_{i,t} \le \bar{\theta}_{i,t} \le C^c$, where $\bar{\theta}_{i,t}$ is an upper bound on $\theta_{i,t}$. Thus, we linearize $w_{i,t}$ as follows:

$$0 \le w_{i,t} \le \theta_{i,t},\;\; \theta_{i,t} - (1-z_{i,t})\bar{\theta}_{i,t} \le w_{i,t} \le \bar{\theta}_{i,t} z_{i,t}. \tag{3.7}$$

3.3.2 Piecewise linearization

We note that for a generator i, its probability of failure by time t, $P_{i,t}$, and maintenance cost at time t, $\theta_{i,t}$, depend nonlinearly on its effective age $\tau_{i,t}$. To accurately capture these nonlinear relationships, we propose linearization procedures for representing $P_{i,t}$ and $\theta_{i,t}$ as functions of $\tau_{i,t}$. For this purpose, we examine the maintenance cost function and the remaining lifetime distribution of each generator under specific breakpoints, namely $d_i^0, d_i^1, \cdots, d_i^K$ for every generator i. Then, we find the corresponding failure probabilities and dynamic maintenance cost function values evaluated at the breakpoints as $P_i^k = \Pr(R_{i,\tau_{i,0}} \le d_i^k)$ and $\theta_i^k = C_{i,\tau_{i,0}}(d_i^k)$ for $k = 1, \cdots, K$, respectively. We illustrate the nonlinearity of the maintenance cost function and its associated piecewise approximation on a sample signal in Figure 3.3. Since we have monthly maintenance decisions, we utilize monthly breakpoints as shown. As remaining lifetime distribution and maintenance cost functions are not convex, we need *special ordered sets of type 2 (SOS2)* constraints in the piecewise linearization, which are a form of disjunctive constraints, see [45]. To formulate these constraints, we consider two formulations studied in [46] with linearly or logarithmically many extra binary variables in the number of breakpoints and constraints. We refer to the first case as *linear* formulation, and second one as *log* formulation. Our preliminary computational results illustrate the significant computational advantage of the log formulation over the linear formulation (see Table 3.4). Therefore, we focus on the log formulation in the remainder of the chapter.

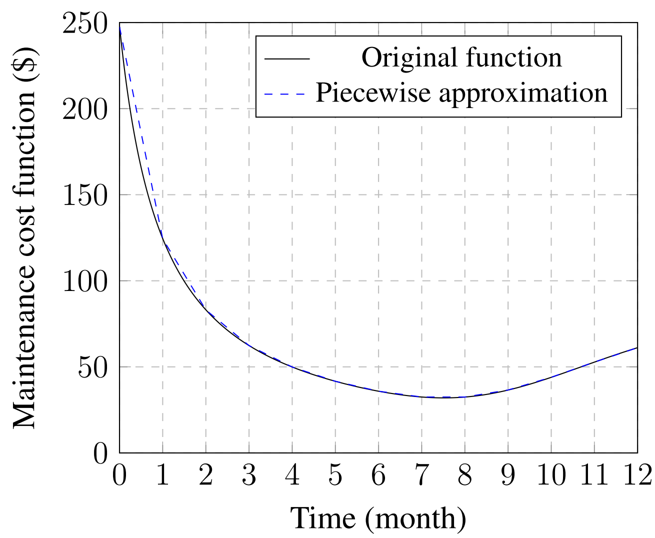


Figure 3.3: Piecewise linearization of maintenance cost function.

The corresponding model can be represented by defining the additional variables $\lambda_{i,t}^k \geq 0$ and $\eta_{i,t}^m \in \{0,1\}$, where $\sum_{k=0}^K \lambda_{i,t}^k = 1$ for all $i \in \mathcal{G}$, $t \in \mathcal{T}$, $k = 0, 1, \cdots, K$, $m \in \mathcal{M}$. The variable $\eta_{i,t}^m$ depends on $\lambda_{i,t}^k$ as follows:

Theorem 1. *(Theorem 1 in [45]) Let $B : \{1, \cdots, K\} \to \{0,1\}^{\lceil \log_2 K \rceil}$ be an SOS2 compatible function, i.e. a function that enforces SOS2 constraints on $\{\lambda_{i,t}^k\}_{k=0}^K \in \mathbb{R}_+^{K+1}$ if for all $l \in \{1, \cdots, K-1\}$ the vectors $B(l)$ and $B(l+1)$ differ in at most one component.*

Then the following inequalities are valid for SOS2 constraints:

$$\sum_{k \in K^+(m,B)} \lambda_{i,t}^k \leq \eta_{i,t}^m \quad m \in \mathcal{M},\ i \in \mathcal{G},\ t \in \mathcal{T}, \tag{3.8a}$$

$$\sum_{k \in K^0(m,B)} \lambda_{i,t}^k \leq (1 - \eta_{i,t}^m) \quad m \in \mathcal{M},\ i \in \mathcal{G},\ t \in \mathcal{T}, \tag{3.8b}$$

where $\mathcal{M} = \{1, \cdots, \lceil \log_2 K \rceil\}$, $K^+(m,B) = \{j \in J : \forall i \in I(j)\ \ m \in \sigma(B(i))\}$, *and* $K^0(m,B) = \{j \in J : \forall i \in I(j)\ \ m \notin \sigma(B(i))\}$. *The function* $\sigma(r)$ *represents the support of vector* r*, which corresponds to the set of indices of* r *such that* $r_i \neq 0$*, and the sets* $I = \{1, \cdots, K\}$, $J = \{0, 1, \cdots, K\}$, $S_i = \{i-1, i\}$ *for all* $i \in I$*, and* $I(j) = \{i \in I : j \in S_i\}$ *for all* $j \in J$.

We note that we use Gray code [47] as the SOS2 compatible function in our formulation, which is used in binary numeral systems to order numbers in such a way that a pair of successive numbers are only different in one binary digit. The remaining constraints, in addition to (3.8), for the piecewise linearization can be expressed as follows:

$$\sum_{k=0}^{K} \lambda_{i,t}^k \theta_i^k = \theta_{i,t} \qquad i \in \mathcal{G},\ t \in \mathcal{T} \tag{3.9a}$$

$$\sum_{k=0}^{K} \lambda_{i,t}^k P_i^k = P_{i,t} \qquad i \in \mathcal{G},\ t \in \mathcal{T} \tag{3.9b}$$

$$\sum_{k=0}^{K} \lambda_{i,t}^k d_i^k = \tau_{i,t} \qquad i \in \mathcal{G},\ t \in \mathcal{T} \tag{3.9c}$$

$$\sum_{k=0}^{K} \lambda_{i,t}^k = 1 \qquad i \in \mathcal{G},\ t \in \mathcal{T} \tag{3.9d}$$

$$\lambda_{i,t}^k \geq 0 \quad k \in \{0, 1, \cdots, K\}, \quad \eta_{i,t}^m \in \{0,1\} \qquad m \in \mathcal{M},\ i \in \mathcal{G},\ t \in \mathcal{T}. \tag{3.9e}$$

3.3.3 Formulation enhancements

We improve the resulting formulation by benefiting from the underlying structure of the problem. The constraint (3.4e) ensures that each generator enters maintenance once during the planning horizon. Therefore, we propose an alternative effective time definition by only considering the values of $\tau_{i,t}$ variables at the time the generators enter maintenance. This information is sufficient for our formulation to compute the maintenance cost function, and the failure probabilities in the chance constraint. Let $\tau'_{i,t} = \tau_{i,t}\, z_{i,t}$. To incorporate this variable in formulation (3.8) and (3.9), we revise the constraints (3.9c) and (3.9d) as follows:

$$\sum_{k=0}^{K} \lambda_{i,t}^{k} d_i^k = \tau'_{i,t} \quad i \in \mathcal{G},\, t \in \mathcal{T} \tag{3.10a}$$

$$\sum_{k=0}^{K} \lambda_{i,t}^{k} = z_{i,t} \quad i \in \mathcal{G},\, t \in \mathcal{T} \tag{3.10b}$$

$$\eta_{i,t}^{m} \le z_{i,t} \quad m \in \mathcal{M},\, i \in \mathcal{G},\, t \in \mathcal{T}. \tag{3.10c}$$

When $z_{i,t}$ is 0, then $\lambda_{i,t}^k$ and $\eta_{i,t}^m$ values are set to 0 for all $k = 0, \cdots, K$ and $m \in \mathcal{M}$, because of (3.10b) and (3.10c), respectively. Thus, $\tau'_{i,t}$ becomes 0 as desired.

This approach provides a significant computational advantage by eliminating the consideration of $\tau_{i,t}$ values when $z_{i,t} = 0$. Similar to the previous linearizations, (3.6) and (3.7), the constraint set $\tau'_{i,t} = \tau_{i,t}\, z_{i,t}$ is linearized, by defining the upper bound value of $\tau'_{i,t}$ as $\overline{\tau}_{i,t}$.

Combining the above, the resulting mathematical problem for the load-dependent maintenance and optimization scheduling (3.4) is reformulated in (3.11) as a mixed-integer linear program. We remove the decision variable $\tau_{i,t}$ from the model as it is no longer needed explicitly.

$$\min\ \xi \left(\sum_{i\in\mathcal{G}} \sum_{t\in\mathcal{T}} w_{i,t} + \sum_{t\in\mathcal{T}} C_{add}\gamma_t \right) + \sum_{i\in\mathcal{G}} \sum_{t\in\mathcal{T}} \sum_{s\in\mathcal{S}} (V_{i,t,s}x_{i,t,s} + F_{i,t,s}y_{i,t,s}) + \sum_{t\in\mathcal{T}} \sum_{s\in\mathcal{S}} \sum_{b\in\mathcal{B}} P_{DC}\psi_{b,t,s} \tag{3.11a}$$

$$\text{s.t. } (3.4d) - (3.4j), (3.6) - (3.8), (3.9a), (3.9b), (3.10)$$

$$\tau'_{i,t} \leq \overline{\tau}_{i,t} z_{i,t} \qquad i \in \mathcal{G}, t \in \mathcal{T} \tag{3.11b}$$

$$\tau'_{i,t} \leq \sum_{t'=1}^{t} \sum_{s \in \mathcal{S}} (\alpha_i y_{i,t',s} + \beta_i x_{i,t',s}) \qquad i \in \mathcal{G}, t \in \mathcal{T} \tag{3.11c}$$

$$\tau'_{i,t} \geq \sum_{t'=1}^{t} \sum_{s \in \mathcal{S}} (\alpha_i y_{i,t',s} + \beta_i x_{i,t',s}) - (1 - z_{i,t})\overline{\tau}_{i,t} \qquad i \in \mathcal{G}, t \in \mathcal{T} \tag{3.11d}$$

$$\tau'_{i,t} \geq 0, \lambda^k_{i,t} \geq 0, \;\; k = 0, 1, \cdots, K, \; i \in \mathcal{G}, \; t \in \mathcal{T}, \;\; \eta^m_{i,t} \in \{0, 1\} \;\; m \in \mathcal{M}, \; i \in \mathcal{G}, \; t \in \mathcal{T} \tag{3.11e}$$

In order to improve the upper bound values used in the linearization, we consider the effect of the load decisions on the data-driven degration equivalent time model. In any time t, the corresponding effective time for generator i, i.e. $\tau_{i,t}$, can be at most $\sum_{t'=1}^{t} \sum_{s \in \mathcal{S}} (\alpha_i p_i^{max} + \beta_i)$, as ensured by the constraint (3.11c). Thus, we can select

$$\overline{\tau}_{i,t} := t\,|\mathcal{S}|(\alpha_i p_i^{max} + \beta_i). \tag{3.12}$$

Similarly, we can identify upper bounds for the failure probabilities $P_{i,t}$ and the dynamic maintenance cost function $\theta_{i,t}$ for generator i at time t. Since $P_{i,t} = \Pr(R_{i,\tau_{i,0}} \leq \tau_{i,t})$ is monotonically nondecreasing with respect to degradation amount, we can take its upper bound value as $\bar{P}_{i,t} = \Pr(R_{i,\tau_{i,0}} \leq \overline{\tau}_{i,t})$. Finally, we can obtain an upper bound value for the dynamic maintenance cost function $\theta_{i,t}$ as $\bar{\theta}_{i,t} = \max_{0 \leq t' \leq \bar{\tau}_{i,t}} C_{i,\tau_{i,0}}(t')$. We note that we are not able to simply select $C_{i,\tau_{i,0}}(\overline{\tau}_{i,t})$ as the upper bound value, since the cost function is not necessarily monotonic with respect to effective time.

3.4 Computational Results

In this section, we provide a comprehensive framework to illustrate the effectiveness of our approach. We first discuss the experimental setup to estimate remaining lifetime and effective time of each generator in Section 3.4.1 and Section 3.4.2, respectively. To evaluate the performances of different maintenance and operations schedules, we develop a decision-dependent simulation procedure in Section 3.4.3. We provide our computational experiments in Section 3.4.4 by studying various instances under different congestion levels and reliability considerations. Finally, we illustrate the computational gains of the proposed algorithmic enhancements for solving the optimization model in Section 3.4.5.

3.4.1 Determining prior distribution and remaining lifetime estimation

To estimate prior distributions corresponding to signal characteristics in (3.1), we first construct a set of 100 signals under different load levels. These signals mimic the degradation process of a rotating bearing, and follow the form (3.1) with $\theta_i \sim N(20, 3^2)$, $\nu_i \sim N(2.5, 0.2^2)$ and $\sigma = 3.5$. We examine the signal values at discrete time points, and assume that load level remains constant between consecutive observations. We observe the signals until a failure threshold Λ, which is taken as 150. Time of failure of each signal i is denoted as T_i. As before, we represent the differences in the observations of each signal i as S_k^i, where $S_1^i = S_i(0)$, and $S_k^i = S_i(k) - S_i(k-1)$ for $k = 2, \cdots, T_i$. Similarly, $L_{i,j}$ corresponds to the load level of signal i in period j. We assume that the variance of the stochastic parameters are known. Thus, we only need to estimate the mean of the prior distributions of the stochastic model parameters θ and ν, which are μ_0 and μ_1 respectively. As the initial amplitude of each signal i, i.e. S_1^i, corresponds to θ_i values in the form (3.1), we compute the mean estimate of $\pi_1(\theta)$ by averaging these values over the set of signals. To estimate the mean of the prior distribution of ν, we find the mean estimates, $\hat{\mu_i}$ corresponding to each signal i, $i = 1, \cdots, 100$. For finding these estimates for the load-dependent models, we adopt the time transformation concept discussed in Section 3.2.2. In particular, we can estimate $\hat{\mu_i}$ as

$$\hat{\mu_i} = \frac{\sum_{j=1}^{T_i} S_j^i - S_1^i}{\sum_{j=1}^{T_i} L_{i,j}}, \tag{3.13}$$

for $i = 1, \cdots, 100$, where the denominator corresponds to an effective time estimate at failure. By averaging these estimates, we obtain the mean estimate of the prior distribution $\pi_2(\nu)$. As the load-independent models neglect the load decisions in determining degradation amount, they consider a different estimate for the prior distribution $\pi_2(\nu)$. Specifically, mean estimate of each signal can be computed as $(\sum_{j=1}^{T_i} S_j^i - S_1^i)/T_i$ by only considering the operational time until failure. Consequently, we obtain the prior estimate for the load-independent model by averaging these values over 100 signals.

In order to represent the signal characteristics specific to each generator, we combine the prior distributions with sensor information. For this purpose, we assign signals to each generator $i \in \mathcal{I}$, which are different than the 100 signals used in prior estimation. These signals are partially degraded at the beginning of the planning horizon with a random initial age. Using these observations until the time of planning, we obtain the posterior distribution of the unknown parameters as discussed in Proposition 3. After obtaining these component specific estimates, we identify the remaining lifetime distribution corresponding to each generator to be used in the optimization model.

3.4.2 Effective time estimation

We utilize the effective time estimation procedure described in Section 3.2.3 using the signals for estimating prior distribution parameters. We consider 3 levels of load, i.e. L_1, L_2, L_3, which corresponds to values 0.5, 1.0, 1.5 respectively. Then, for each signal and load level, we find the estimates using (3.2). Therefore, we obtain the set of points $(L_j, \overline{\mu}_{i,j})$ for each load level $j \in \{1, 2, 3\}$ and signal $i \in \{1, \cdots, 100\}$. These points for the given signal set are illustrated in Figure 3.4, which shows the variability in the signals under each load level. We note that the overall drift average over the signals, $\overline{\overline{\mu}}$, is computed as 2.49.

For every generator i, we consider the values p_i^{min}, $(p_i^{min} + p_i^{max})/2$, and p_i^{max} corresponding to the load levels L_1, L_2, L_3, respectively. Next, we rescale the drift parameter estimates by dividing them to the population mean estimate $\overline{\overline{\mu}}$. Consequently, the rescaled vertical axis represents the unit effective time. By applying regression analysis specific to each generator i, we find the estimates α_i and β_i to be used in the optimization model (3.11).

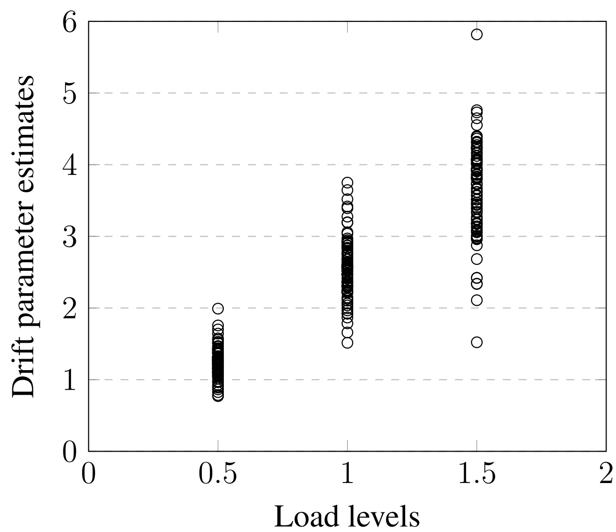


Figure 3.4: Signal variability under each load level.

We remind the reader that load-independent models disregard the notion of effective time based on the production decisions. To capture this approach, we replace constraint (3.4b) with

$$\tau_{i,t} = \sum_{t'=1}^{t} \sum_{s \in \mathcal{S}} \frac{1}{|\mathcal{S}|} x_{i,t',s} \quad i \in \mathcal{G}, t \in \mathcal{T}, \tag{3.14}$$

which gives the operational age of the generator. Then, we proceed with the same formulation methodology as in the load-dependent case in Section 3.3.2 and Section 3.3.3 to linearize and represent the variable $\tau'_{i,t}$. Consequently, effective time in these models are only based on the operational age of the generators. By coupling this condition with the remaining lifetime estimations specific to the load-independent models, we can model the maintenance and operations scheduling problem under solely operational time-based degradation.

3.4.3 Simulation and solution evaluation

In order to compare the performances of different maintenance and operations schedules, we propose a decision-dependent simulation procedure. In each period, we simulate the degradation process of each generator by creating signals based on the signal characteristics, and the dispatch and commitment decisions. For simulation purposes, we assume that the true distribution of θ_i and μ_i values in the functional form (3.1) are known for each generator $i \in G$. For representing effective time, we use α_i, β_i values found in Section 3.4.2.

Algorithm 4 describes the proposed methodology to evaluate a given maintenance and operations schedule in detail. We start the simulation procedure by considering the last observed signal amplitude of each generator at the beginning of planning. Since we observe each signal i until time t_k^i, we represent the last amplitude as $\sum_{j=1}^{k^i} S_j^i$ as discussed in Section 3.2.2. Then, we simulate each generator's corresponding signal under the given operations schedule. At the end of each maintenance period, we check the condition of the generators by observing their signal amplitudes. If the period t is the scheduled maintenance time of generator i, i.e. $z_{i,t} = 1$, and the generator has not failed previously, then generator enters preventive maintenance and remains closed for Y_p periods. If the signal amplitude of generator i, namely Amp_i, is greater than the failure threshold Λ, then generator i fails. It enters corrective maintenance immediately and stays closed for Y_c periods. As failures are unexpected, corrective maintenance requires more resources than a scheduled maintenance, i.e. $Y_c > Y_p$. After a maintenance ends, a new signal is assigned to that generator to represent its degradation process in the remainder of the planning horizon. Since components start degrading after their first phase ends, the degradation process starts after maintenance is completed, and first phase is over.

Algorithm 4 Solution Evaluation

Obtain z^*, y^*, x^*, ψ^* solutions from the optimization model (3.11).

Set $numPaths = 1000$, $numFailures = 0$, $maintCost = 0$, $totalCost = 0$, $FP = firstPhase$.

for all $l \in \{1, \cdots, numPaths\}$ **do**

  for all $i \in G$ **do**

    $Amp_i = \sum_{j=1}^{k^i} S_j^i$, $hasMainted = False$, $hasFailed$ = False, $maintCompPeriod = 0$.

Generate initial amplitude after maintenance, $newAmp_i$, from the distribution of θ_i.

for all $t \in \{1, \cdots, H\}$ **do**

if $z^*_{i,t} == 1$ and $hasFailed$ == False **then**

$Amp_i = newAmp_i$, $maintCost += C^p$, $hasMainted$= True, $maintCompPeriod = t + Y_p$.

else if $Amp_i > \Lambda$ **then**

$Amp_i = newAmp_i$, $numFailures += 1$, $maintCost += C^c$, $hasMainted = True$, $hasFailed$ = True, $maintCompPeriod = t + Y_c$.

else if $hasMainted$ == False or ($t \geq maintCompPeriod + FP$ and $hasMainted == True$) **then**

Calculate unit degradation d in period t by $d = \sum_{s\in\mathcal{S}}(\alpha_i y^*_{i,t,s} + \beta_i x^*_{i,t,s})$.

$Amp_i += \mu_i d + \sigma m d$, where m is sampled from $N(0,1)$.

end if

if $t \geq maintCompPeriod$ **then**

$totalCost += \sum_{s\in\mathcal{S}}(V_{i,t,s} y^*_{i,t,s} + F_{i,t,s} x^*_{i,t,s})$.

else

$totalCost += \sum_{s\in\mathcal{S}} P_{DC} y^*_{i,t,s}$.

end if

end for

end for

$totalCost += \sum_{t\in\mathcal{T}}\sum_{s\in\mathcal{S}}\sum_{b\in\mathcal{B}} P_{DC}\psi^*_{b,t,s}$.

end for

$totalCost += maintCost$.

Divide $numFailures$, $maintCost$, $totalCost$ by $numPaths$ to find the mean results.

When a generator fails unexpectedly, it will not be able to produce in the upcoming Y_c periods. This unexpected loss in the production needs to be explicitly taken into account while evaluating the maintenance schedule. For this purpose, we consider these types of losses in production due to failures as demand curtailment while computing the operational cost in simulation.

3.4.4 Computational experiments

In this section, we present a comprehensive computational study by comparing the performance of the solutions from load-dependent and load-independent models. We evaluate these solutions using the simulation procedure described in Section 3.4.3. We provide our computational results on 39-bus New-England Power System [23], and 118-bus instances [24]. An overview of the instances is provided in Table 3.1, and further details of the power system configurations are discussed in the aforementioned papers and references therein. We implement the proposed model with enhancements (3.11) in Python using Gurobi 7.5.2 as the solver on an Intel i5-3470T 2.90 GHz machine with 8 GB RAM.

Table 3.1: Overview of the Instances.

	# Buses	# Lines	# Generators	Total capacity (MWh)
39-bus	39	46	10	8840.4
118-bus	118	186	19	5859.2

We study a one-year maintenance plan with monthly maintenance and daily operational decisions. For the chance constraint (3.4c), we set ρ as $\lfloor |\mathcal{G}|/3 \rfloor$ with $\epsilon = 0.05$ or 0.10. This implies that at most one third of the generators enters corrective maintenance due to a failure with a probability of at least $(1-\epsilon)$. The safe approximation discussed in Proposition 4 is used to represent the chance constraint. We set cost of preventive maintenance C^p = \$100.000, and corrective maintenance C^c = \$400.000. These cost values are used in both dynamic maintenance cost function calculation in (3.3), and in the simulation for evaluating maintenance costs. To observe the performance of the proposed approach under various signal characteristics, we generate a partially degraded set of signals following the procedure in Section 3.4.1. Then, we randomly assign these signals to the generators and repeat each experiment 5 times with different set of signals. For each setting, we report the average results of these 5 macro-replications.

We evaluate three modeling approaches with respect to their remaining lifetime estimation procedures and optimization formulations, namely i) load-dependent, ii) load-independent, and iii) reliability-based. Load-dependent refers to the proposed approach of the chapter to represent the decision-dependent degradation in maintenance and operations scheduling. Load-independent and reliability-based approaches consider an operational age-based degradation modeling as represented in (3.14) in the optimization model (3.11), whereas they differ in their remaining lifetime estimation procedures. Load-independent approach adopts its estimation procedure described in Section 3.2.2. For the reliability-based case, we derive the lifetime distributions by first fitting an inverse Gaussian distribution to a given set of failure points of the signals used in prior estimation, and then conditioning to the initial ages of the generators.

Table 3.2: Solution Evaluation under High Congestion.

	ϵ	Type	# of Failures	MC ($M)	Gain (%)	TC ($M)	Gain (%)
39-bus	0.05	LD	0.42	1.13		113.30	
		LI	1.31	1.39	18.74	129.62	12.59
		RB	0.63	1.19	4.94	117.62	3.68
	0.10	LD	0.26	1.08		111.14	
		LI	1.18	1.35	20.24	124.84	10.97
		RB	1.04	1.32	17.91	124.13	10.47
118-bus	0.05	LD	0.66	2.10		74.00	
		LI	2.11	2.53	16.97	80.79	8.41
		RB	1.64	2.40	12.25	80.20	7.73
	0.10	LD	0.60	2.08		72.24	
		LI	2.14	2.54	18.09	81.42	11.27
		RB	1.63	2.39	12.89	80.58	10.35

As demand level of the system plays an important role in the maintenance and operations decisions and the degradation amount of the generators, we study the instances under two congestion levels. Table 3.2 and Table 3.3 correspond to the results under high and low system congestions, respectively. In particular, low and high congestion correspond to the cases where the average daily demand of the system over a yearly planning horizon is adjusted to be 40% and 70% of the system capacity. The columns '# of failures', ‘MC’ and ‘TC’ represent the average number of failures, maintenance and total cost (sum of maintenance and operations costs) in the evaluated simulation procedure for a given solution. Each instance is studied under two different reliability levels of chance constraint, which is adjusted by the parameter ϵ. The abbreviations ‘LD’, ‘LI’, ‘RB’ are used for load-dependent, load-independent and reliability-based approaches, respectively.

Table 3.3: Solution Evaluation under Low Congestion.

	ϵ	Type	# of Failures	MC ($M)	Gain (%)	TC ($M)	Gain (%)
39-bus	0.05	LD	0.04	1.01		53.02	
		LI	0.60	1.18	14.12	59.82	11.36
		RB	0.10	1.03	1.79	53.83	1.51
	0.10	LD	0.06	1.02		53.14	
		LI	0.65	1.20	14.82	60.08	11.54
		RB	0.12	1.04	1.79	53.75	1.13
118-bus	0.05	LD	0.08	1.93		37.30	
		LI	0.94	2.18	11.58	40.70	8.35
		RB	0.49	2.05	5.87	38.33	2.68
	0.10	LD	0.05	1.92		37.31	
		LI	1.06	2.22	13.49	41.46	10.01
		RB	0.51	2.05	6.55	38.31	2.62

As the proposed load-dependent approach captures the effect of operational decisions on degradation modeling within the optimization model, it performs better in terms of number of failures and maintenance cost in comparison to load-independent and reliability-based approaches. We observe 5-20% and 2-15% maintenance cost savings of load-dependent approach in high and low system congestions, respectively, compared to the previously studied methods in the literature.

When a generator fails unexpectedly, there is an unplanned loss in production capacity. This disruption in the operational schedule is penalized with demand curtailment cost in the solution evaluation. Our analyses highlight significant cost savings in total cost in the order of 3-13% and 1-11% for high and low congestion cases, by adopting the load-dependent approach. Furthermore, when systems are under high congestion, we observe more failures in all instances. This happens since high demand levels initiate higher levels of production resulting in faster degradation.

We emphasize that load-independent and reliability-based approaches are insufficient in truly representing the dependency between the degradation modeling and optimization framework. Nevertheless, as we compare the two approaches, we observe that reliability-based schedules perform better in solution evaluation compared to load-independent schedules. Reliability-based remaining lifetime estimations consider more variance in data by fitting a lifetime distribution based on failure points, whereas load-independent estimations are tailored to unit specific observations. Consequently, maintenance cost function and remaining lifetime estimations of reliability-based approach do not change much between different maintenance decisions, compared to load-independent models. Thus, the resulting optimization model becomes less sensitive to the choice of maintenance and operations schedules.

We note that as ϵ value gets larger, the chance constraint becomes less restrictive. Although the choice of ϵ does not necessarily affect the performance of the resulting schedule in high congestion case, average number of failures decreases when $\epsilon = 0.05$ in low congestion setting. Overall, the results demonstrate that load-dependent solutions outperform load-independent and reliability-based solutions with a smaller number of failures, and lower maintenance and operational costs in all the settings considered.

3.4.5 Computational efficiency

In this section, we illustrate the computational efficiency with respect to different forms of enhancements. We first examine the computational advantage of the selected piecewise linearization procedure, used for linearizing the objective and the safe approximation of the chance constraint. Specifically, we compare linear and log formulations described in Section 3.3.2. Secondly, we illustrate the performance of the proposed formulation enhancements in Section 3.3.3. Lastly, we demonstrate the effect of a *priority branching* method. More specifically, as maintenance decisions play an important role in determining the effective time, we put a special emphasis on those variables while solving the problem. For this purpose, we employ a priority branching method, which is used in optimization for directing the branch-and-bound procedure. By prioritizing the variables corresponding to the maintenance decisions, z, over the commitment decisions x, the respective branch-and-bound tree prefers branching on the maintenance variables.

We demonstrate the results on sample 39-bus instances under high congestion and $\epsilon = 0.05$ in Table 3.4. We report the average run time results over 5 macro-replications. We note that the linear formulation is not able to converge in 10000 seconds in all replications. By log formulation, we refer to the proposed integrated maintenance and operations scheduling model without the formulation enhancements introduced in Section 3.3.3.

Table 3.4: Run time (seconds) comparison on sample instances.

	Run time	Speed-up
Linear formulation	$>$10000.00	
Log formulation	1189.89	$\times$8.40
Log formulation with enhancements (3.11)	601.53	$\times$16.62
Log formulation with enhancements (3.11) + priority branching	576.41	$\times$17.35

The results show that each enhancement significantly contributes to the run time performance. Priority branching coupled with the formulation enhancements gives the best results with an overall speedup of more than 17 times, demonstrating the significant computational gains of the proposed improvements.

3.5 Concluding Remarks

In this study, we present a comprehensive framework for effectively solving condition-based maintenance and operations scheduling problem of a fleet of generators under load-dependency. We propose a data-driven degradation modeling framework to capture the endogenous effect of the operational decisions. First, we present a sensor-driven remaining lifetime estimation procedure under time-varying load decisions. We also develop an estimation method to capture the effect of the load decisions while taking into account the signal variability. We formulate a novel stochastic optimization model and propose a piecewise linearization method for accurately representing the operational decisions' effect on the degradation models in combination with other formulation enhancements. We also extend the chance constraint proposed in Chapter 1 to the decision-dependent setting. To evaluate the performances of the maintenance schedules, we develop a decision-dependent simulation framework. This framework enables determining the quality of a solution by simulating signals based on a given schedule. We provide a comprehensive computational study on two illustrative IEEE test cases by comparing the proposed load-dependent approach with load-independent and reliability-based approaches. We present a computational analysis by optimizing the schedules under different congestion levels and conservativeness amount of the chance constraint. Our analysis demonstrates the superior performances of the load-dependent schedules with reductions in failures and significant cost savings up to 20%. Finally, we provide experiments demonstrating the computational efficiency of the formulation improvements up to 17 times speedup. These results highlight the importance of considering operational decisions in condition-based maintenance scheduling to ensure reliableness and cost effectiveness of the system.

CHAPTER 4

ADAPTIVE TWO-STAGE STOCHASTIC PROGRAMMING WITH AN APPLICATION TO CAPACITY EXPANSION PLANNING

4.1 Introduction

Optimization in sequential decision making processes under uncertainty is known to be a challenging task. Two-stage and multi-stage stochastic programming are fundamental techniques for modeling these processes, where stage refers to the decision times in planning. In two-stage programs, a set of decisions need to be determined at the beginning of the planning horizon resulting in *static* policies, whereas multi-stage programs allow total flexibility by deriving *fully adaptive* policies depending on the observed uncertainty. Although both approaches have its own pros and cons, the resulting policies may not be sufficient to address a wide range of business settings due to the flexibility level of the corresponding processes. Specifically, one may need to have a fixed order of decisions before and after a specified time period as detailed below. To address these issues, we propose a *partially adaptive* stochastic programming approach that determines the best time to revise the decisions for the problems with limited flexibility.

There have been many problems in the literature that require partially adaptive policies for determining the best set of actions over a multi-period planning horizon. We will now motivate the applicability of partially adaptive approaches for three specific example settings:

- Capacity expansion management is a strategic level planning problem to determine the expansion times and amount of different resources in areas such as electricity expansion, production planning and network design. This problem involves uncertainties in system demand and investment costs. Setting the expansion decisions at the beginning of the planning results in static and restrictive policies, corresponding to two-stage models. Nevertheless, these actions might not be updated in each period as in multi-stage models since the expansion of resources may require commitments and lead time for establishing the necessary infrastructure. Therefore, fully adaptive policies obtained from multi-stage models may not be feasible either.

- Another example setting for the partially adaptive approaches involves portfolio optimization problems. To construct a portfolio, one may need to determine a fixed sequence of investment decisions for a period of time with a possible option to revise in future. Two-stage stochastic programs may not be sufficient as they entail rigid schedules by not allowing any revision option at all. On the other hand, it may not be appropriate to update decisions in each period due to additional transaction costs associated with rebalancing actions.

- A similar problem occurs in major overhaul decisions of components in maintenance scheduling. The overhaul schedules need to be determined ahead over a multi-year plan with a possibility of revision depending on the changes in components' conditions over time. Two-stage approaches result in restrictive schedules by not considering any change at all. On the other hand, it might be difficult or costly to observe the system's situation and update the schedules at each time period. Therefore, multi-stage approaches may not be applicable for this setting either.

As it can be observed from these three applications, static approaches may not be sufficient to address these problems by not allowing any revision in the schedules. Similarly, fully adaptive approaches may not be suitable either due to the problem characteristics and difficulty of obtaining system's state at each period. Therefore, in this chapter, we will focus on the analysis of an adaptive two-stage approach by optimizing the revision points of each decision.

We first review the relevant studies in the literature that adopt partially adaptive policies, specifically developed for inventory and lot sizing problems. These problems focus on determining the inventory and production decisions under nonstationary demand and cost structures over a multi-period planning horizon. Partially adaptive policies are motivated in these problem settings for establishing the coordination and synchronization of the supply chain systems over dynamic policies [48]. [49] introduces the *static-dynamic* uncertainty strategy for solving a probabilistic lot sizing problem by selecting the replenishment times at the beginning of the planning horizon and determining the corresponding order quantities at these time points. [50] extends this concept by developing a mixed-integer linear programming formulation. Variants of this strategy have been studied in [51, 52, 53, 54] to address the stochastic lot sizing problem under different demand and cost functions, and service level constraints.

In the capacity expansion planning literature, uncertainties over a multi-period planning horizon are addressed with different methods (see [55] for an extensive survey). [56] studies capacity acquisition decisions and their timing to meet customer demand according to the technological breakthroughs by adopting a dynamic programming based solution methodology. Similarly, stochastic dynamic programming has been applied to this problem [6, 57], despite its disadvantages in incorporating the practical constraints. Stochastic programming is another fundamental methodology to address these problems by representing the underlying uncertainty through scenarios. For instance, [58] and [59] model this problem as a two-stage stochastic program by first determining the capacity expansion decisions and then adapting the capacity allocations with respect to the scenarios. On the other hand, [60], [61], and [62] consider these problems as multi-stage stochastic mixed-integer programs and represent uncertainties through scenario trees. Despite this extensive literature, these studies neglect the need for partial flexibility in capacity expansion problems.

In the stochastic programming literature, intermediate approaches between two-stage and multi-stage models have been studied under different problem contexts to mainly address the computational complexity associated with the multi-stage models. As an example, *shrinking-horizon strategy* involves solving two-stage stochastic programs between predetermined time windows. Specifically, [63] and [64] consider this strategy to obtain an approximation to multi-period planning problems in oil industry, and multi-product batch plant under demand uncertainty for chemical processes, respectively. Shrinking-horizon strategy is also applied to airline revenue management problem in [65] by proposing heuristics to determine the resolve points under specific assumptions regarding the stochastic process. Another two-stage approximation to multi-stage models is presented in [66] using Linear Decision Rules by limiting the decisions to be affine functions of the uncertain parameters. As an alternative intermediate approach, [67] proposes solving a generation capacity expansion planning problem by first considering a multi-stage stochastic program until a predefined stage, and then representing it as a two-stage program. They also develop a *rolling horizon heuristic* as discussed in [68] to approximate the multi-stage model. Another line of research [69] addresses how many stages to have in a multi-stage stochastic program by contamination technique which focuses on limiting the deviations from the underlying uncertainty distribution. These results are then extended to problems with polyhedral risk objectives for financial optimization problems in [70]. Several other studies [71, 72] focus on numerical analyses for the choice of planing horizon and stages specifically for portfolio management. However, there has been little to no emphasis on optimizing the stage decisions by determining best time to observe uncertainty for a generic problem setting.

In this study, we propose a partially adaptive stochastic programming approach, in which the revision points are decision variables. Specifically, we consider a fixed sequence of decisions before and after a specific revision point, which is optimized for each decision. This procedure provides significant advantages for the settings where partially adaptive approaches become necessary, and two- and multi-stage models are not appropriate. Our contributions can be summarized as follows:

1. We propose an adaptive two-stage stochastic programming approach, in which we

optimize the revision decisions over a static policy. We analyze this approach using a multi-period newsvendor problem, and provide a policy under the adaptive setting. We then develop a mixed-integer linear programming formulation for representing the proposed approach, and prove the NP-Hardness of the resulting stochastic program.

2. We provide analyses on the value of the proposed approach compared to two-stage and multi-stage stochastic programming methods with respect to the choice of the revision decisions. We focus our analyses on a specific structure that encompasses capacity expansion planning problem.

3. We propose solution algorithms for the adaptive two-stage program under the studied problem structure. We provide an approximation guarantee and demonstrate its asymptotic convergence.

4. We demonstrate the benefits of the adaptive two-stage approach on a generation capacity expansion planning problem. Our extensive computational study illustrates the relative gain of the proposed approach with up to 21% reduction in cost, compared to two-stage stochastic programming over different scenario tree structures. Our results also highlight the significant run time improvements in approximating the desired problem with the help of proposed solution algorithms. We also analyze a sample generation expansion plan to examine the practical implications of optimizing revision decisions.

The remainder of this chapter is organized as follows: In Section 4.2, we motivate the adaptive two-stage approach using the newsvendor problem and provide an analytical analysis for its optimal policy. In Section 4.3, we formally introduce the adaptive two-stage stochastic programming model. In Section 4.4, we study the proposed approach on a class of problems encompassing the capacity planning problem and present analytical results on its performance in comparison to the existing methodologies. In Section 4.5, we develop solution methodologies and derive their approximation guarantees by benefiting from our analytical results. In Section 4.6, we present an extensive computational study on a generation expansion planning problem in power systems. Section 4.7 concludes the chapter with final remarks and future research directions.

4.2 A motivating example: The newsvendor problem

4.2.1 Formulation and optimal policy

In this section, we illustrate the adaptive two-stage approach on a newsvendor problem with T periods. The decision maker determines the order amount in each period t, namely x_t, while minimizing the total expected cost over the planning horizon. Demand in period t, denoted by d_t for $t = 1, \cdots, T$, is assumed to be random and independently distributed across periods. We consider a unit holding cost at the end of each period t as h_t, and assume that stockouts are backordered with a cost b_t. We incur an ordering cost c_t per unit in each period t and assume that initial inventory at hand is zero. We also assume that the cost structure satisfies the relationship $c_t - b_t \leq c_{t+1} \leq c_t + h_t$ for $t = 1, \cdots, T-1$, because otherwise it might become more profitable to backorder demand or hold inventory. Additionally, we set $b_T \geq c_T$ to avoid backordering at the end of the planning horizon.

Our goal is to formulate an adaptive program in which we determine the order schedule until a specified time period t^*, then observe the underlying uncertainty and determine the remainder of the planning horizon accordingly. We first note that by defining inventory amount at the end of period t as I_t and considering the inventory relationship $I_t = I_{t-1} + x_t - d_t$, we can rule out the inventory variable using the relationship $I_t := \sum_{t'=1}^{t}(x_{t'} - d_{t'})$. Consequently, dynamic programming formulation of the adaptive multi-period newsvendor problem with a revision point at t^* can be formulated as follows:

$$\min_{x_1,\cdots,x_{t^*-1}} \Big\{ \sum_{t=1}^{t^*-1} \big(c_t x_t + E[h_t \max\{\sum_{t'=1}^{t}(x_{t'} - d_{t'}), 0\} - b_t \min\{\sum_{t'=1}^{t}(x_{t'} - d_{t'}), 0\}]\big) + E[Q_{t^*}(\sum_{t=1}^{t^*-1}(x_t - d_t))]\Big\}, \quad (4.1)$$

where

$$Q_{t^*}(s) = \min_{x_{t^*},\cdots,x_T} \Big\{ \sum_{t=t^*}^{T} \big(c_t x_t + E[h_t \max\{s + \sum_{t'=t^*}^{t}(x_{t'} - d_{t'}), 0\} - b_t \min\{s + \sum_{t'=t^*}^{t}(x_{t'} - d_{t'}), 0\}]\big)\Big\}. \quad (4.2)$$

Theorem 2. *Order quantity for the adaptive two-stage approach* (4.1) *can be represented in the following form:*

$$\widetilde{F}_{1,t}(X_{1,t}) = \frac{-c_t + c_{t+1} + b_t}{h_t + b_t}, \qquad t = 1, \cdots, t^* - 1, \quad (4.3)$$

$$\widetilde{F}_{t^*,t}(s_{t^*} + X_{t^*,t}) = \frac{-c_t + c_{t+1} + b_t}{h_t + b_t}, \qquad t = t^*, \cdots, T, \quad (4.4)$$

where $X_{i,j} = \sum_{t=i}^{j} x_t$*,* $D_{i,j} = \sum_{t=i}^{j} d_t$*,* $s_{t^*} = X_{1,t^*-1} - D_{1,t^*-1}$*,* $c_{T+1} = 0$*, and* $\widetilde{F}_{i,j}$ *is the cumulative distribution function of* $D_{i,j}$*.*

Proof. To obtain an optimal policy for the adaptive two-stage model (4.1), we first focus on the problem (4.2) where the decision maker observes the inventory at hand at the beginning of period t^* to determine the order schedule of periods $t^*, \cdots, T$.

Proposition 5. *For the problem* (4.2)*, we obtain an optimal policy in the following form:*

$$\widetilde{F}_{t^*,t}(s + X_{t^*,t}) = \frac{-c_t + c_{t+1} + b_t}{h_t + b_t}, \quad t = t^*, \cdots, T. \quad (4.5)$$

Proof. We first observe that the objective function of the problem (4.2) is convex as it can be represented as sum of convex functions. Then, we rewrite the objective function in the following form:

$$H(x) = \sum_{t=t^*}^{T} (c_t x_t + h_t \int_0^{s+\sum_{t'=t^*}^{t} x_{t'}} (s + \sum_{t'=t^*}^{t} x_{t'} - z) \tilde{f}_{t^*,t}(z) dz - b_t \int_{s+\sum_{t'=t^*}^{t} x_{t'}}^{\infty} (s + \sum_{t'=t^*}^{t} x_{t'} - z) \tilde{f}_{t^*,t}(z) dz),$$

where $\tilde{f}_{i,j}$ is the convolution probability density function corresponding to $\sum_{t=i}^{j} d_t$. We then take derivative of this expression with respect to the order quantities of the periods $t^*, \cdots, T$ as follows:

$$\frac{\partial H(x)}{\partial x_{t^*}} = c_{t^*} + h_{t^*} \widetilde{F}_{t^*,t^*}(s + x_t^*) - b_{t^*}(1 - \widetilde{F}_{t^*,t^*}(s + x_t^*)) + \cdots + h_T \widetilde{F}_{t^*,T}(s + \sum_{t=t^*}^{T} x_t) - b_T(1 - \widetilde{F}_{t^*,T}(s + \sum_{t=t^*}^{T} x_t))$$

$$\vdots$$

$$\frac{\partial H(x)}{\partial x_{T-1}} = c_{T-1} + h_{T-1} \widetilde{F}_{t^*,T-1}(s + \sum_{t=t^*}^{T-1} x_t) - b_{T-1}(1 - \widetilde{F}_{t^*,T-1}(s + \sum_{t=t^*}^{T-1} x_t)) + h_T \widetilde{F}_{t^*,T}(s + \sum_{t=t^*}^{T} x_t) - b_T(1 - \widetilde{F}_{t^*,T}(s + \sum_{t=t^*}^{T} x_t))$$

$$\frac{\partial H(x)}{\partial x_T} = c_T + h_T \widetilde{F}_{t^*,T}(s + \sum_{t=t^*}^{T} x_t) - b_T(1 - \widetilde{F}_{t^*,T}(s + \sum_{t=t^*}^{T} x_t))$$

By equating the above expressions to 0, we obtain the desired conditions, concluding the proof. □

As a next step, we aim extending our result to the adaptive two-stage case. We observe that the objective function denoted in (4.1) is convex. Let the corresponding objective function be $G(x)$, and we take its derivative with respect to the order quantities of the periods $1, \cdots, t^*-1$ as follows:

$$\frac{\partial G(x)}{\partial x_1} = c_1 + h_1 \widetilde{F}_{1,1}(x_1) - b_1(1 - \widetilde{F}_{1,1}(x_1)) + \cdots$$
$$+ h_{t^*-1}\widetilde{F}_{1,t^*-1}(\sum_{t=1}^{t^*-1} x_t) - b_{t^*-1}(1 - \widetilde{F}_{1,t^*-1}(\sum_{t=1}^{t^*-1} x_t)) + \frac{\partial E[Q_{t^*}(\sum_{t=1}^{t^*-1}(x_t - d_t))]}{\partial x_1}$$
$$\vdots$$
$$\frac{\partial G(x)}{\partial x_{t^*-1}} = c_{t^*-1} + h_{t^*-1}\widetilde{F}_{1,t^*-1}(\sum_{t=1}^{t^*-1} x_t) - b_{t^*-1}(1 - \widetilde{F}_{1,t^*-1}(\sum_{t=t^*}^{t^*-1} x_t)) + \frac{\partial E[Q_{t^*}(\sum_{t=1}^{t^*-1}(x_t - d_t))]}{\partial x_{t^*-1}}$$

To derive $\frac{\partial Q_{t^*}(s)}{\partial x_i}$ for $i = 1, \cdots, t^*-1$, we use the chain rule as $\frac{\partial Q_{t^*}(s)}{\partial x_i} = \frac{\partial Q_{t^*}(s)}{\partial s} \cdot \frac{\partial s}{\partial x_i}$, where $s = \sum_{t=1}^{t^*-1}(x_t - d_t)$. For that purpose, we need to identify the derivative of $Q_{t^*}(s)$ with respect to s.

Lemma 1. $\dfrac{\partial Q_{t^*}(s)}{\partial s} = -c_{t^*}$.

Proof. We let the optimal solution of the problem (4.2) as $(x^*_{t^*}, \cdots, x^*_T)$. Then, we can represent $\dfrac{\partial Q_{t^*}(s)}{\partial s}$ as follows:

$$\frac{\partial Q_{t^*}(s)}{\partial s} = h_{t^*}\widetilde{F}_{t^*,t^*}(s + x^*_{t^*}) - b_{t^*}(1 - \widetilde{F}_{t^*,t^*}(s + x^*_{t^*})) + \cdots$$
$$+ h_T \widetilde{F}_{t^*,T}(s + \sum_{t=t^*}^{T} x^*_t) - b_T(1 - \widetilde{F}_{t^*,T}(s + \sum_{t=t^*}^{T} x^*_t))$$

Using the optimality conditions derived in the proof of Proposition 5, we can prove that $\dfrac{\partial Q_{t^*}(s)}{\partial s} = -c_{t^*}$. □

Combining the above, we obtain the desired result. □

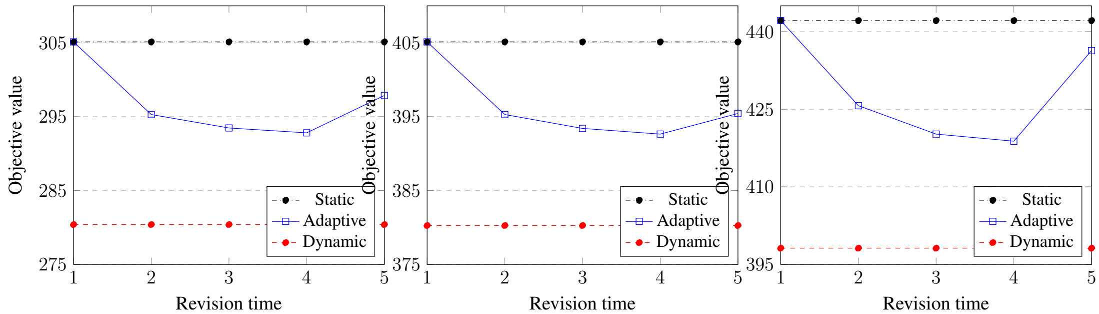
Using Theorem 2, we can show that the optimal adaptive two-stage solution follows an order up-to policy. Let $\{X^*_{1,t}\}_{t=1}^T$ be the cumulative order quantities obtained by Theorem 2. We have i) $X^*_{1,t} = \widetilde{F}^{-1}_{1,t}(\frac{-c_t + c_{t+1} + b_t}{h_t + b_t})$ for $t = 1, \cdots, t^*-1$, and ii) $X^*_{t^*,t} = \widetilde{F}^{-1}_{t^*,t}(\frac{-c_t + c_{t+1} + b_t}{h_t + b_t}) - s_{t^*}$ for $t = t^*, \cdots, T$. Next, we derive the order amount of each period as follows: $x^*_1 = X^*_{1,1}$, and $x^*_t = \max\{X^*_{1,t} - X^*_{1,t-1}, 0\}$ for $t = 2, \cdots, t^*-1$. At time

t^*, we observe the cumulative net inventory of that period, s_{t^*}. Then, we derive the remaining ordering policy as $x^*_{t^*} = \max\{X^*_{t^*,t^*}, 0\}$, and $x^*_t = \max\{X^*_{t^*,t} - \sum_{t'=t^*}^{t-1} x^*_{t'}, 0\}$ for $t = t^* + 1, \cdots, T$. We note than when t^* is set to 1, and s represents the initial inventory, then the adaptive approach converts into a fully static setting where the decision maker determines the order schedules until the end of the planning horizon ahead of the planning.

4.2.2 Illustrative example

To demonstrate the importance of the time to revise our decisions, we illustrate the performance of the adaptive approach under different revision times. We consider $T = 5$, and assume demand in each period is normally and independently distributed with $d_t \sim N(\mu_t = 10, \sigma_t^2 = 4)$, and cost values are set to $c_t = 5$, $h_t = 2$ for $t = 1, \cdots, 5$ in stationary setting. We let $b_t = h_t$ for the first 4 periods, and set $b_5 = c_5 + 1$ to ensure backordering is costly in the last period.

We demonstrate how policies are affected from different revision times by evaluating them under 1000 different demand scenarios in Figure 4.1, and compare them with static and dynamic order up-to policies (see [73]). Specifically, order schedule is determined ahead of the planning in static setting, and ordering decisions can be revised in each period by observing the underlying demand in dynamic policies. We illustrate three cases depending on the cost and demand parameters. As expected, we observe that the fully static and dynamic cases result in the highest and lowest costs, respectively. For the adaptive approach, revising at 4^{th} period gives the least cost in all settings, and objective function value is significantly affected by the choice of the revision time.



(a) Stationary costs, stationary demand.

(b) Stationary costs, increasing demand ($\mu_t = 10 + 2t$).

(c) Increasing costs, ($c_t = 5 + t$, $h_t = b_t = 2 + t$), stationary demand.

Figure 4.1: Objective values under different revision times.

4.3 Adaptive Two-Stage Stochastic Programming Formulation, Complexity and Value

In this section, we first propose a generic formulation of the adaptive two-stage approach. Then, we compare the adaptive two-stage approach with the existing stochastic programming methodologies along with their respective decision structures. Finally, we show that solving the adaptive two-stage stochastic programming is NP-hard.

4.3.1 Generic formulation of adaptive two-stage approach

We first describe a generic formulation for the adaptive two-stage approach, and then extend it to a stochastic setting. We consider a sequential decision making problem with T periods, in which we take into account two sets of decisions. The *state* variables $\{x_t(\xi_{[t]})\}_{t=1}^T$ represent the primary decisions given the data vector $\xi_{[t]}$, namely the data available until period t. These variables are used for linking decisions of different time periods to each other. The *stage* variables $\{y_t(\xi_{[t]})\}_{t=1}^T$ correspond to the secondary decisions that are local to period t. We assume that each state and stage variable have dimensions of I and J, respectively. We formalize the adaptive two-stage approach by allowing one revision decision for each state variable throughout the planing horizon. More specifically, the decision maker determines her decisions for the state variable $x_{it}(\xi_{[t]})$ until its revision time t_i^* for every $i \in \{1, \cdots, I\}$. Then, she observes the underlying data until that period, namely

$\xi_{[t_i^*]}$, and revises the decisions of the corresponding state variable for the remainder of the planning horizon. Combining the above, we formulate the adaptive two-stage program as follows:

$$
\begin{aligned}
\min_{x,\hat{x},y,r} \quad & \sum_{t=1}^{T} f_t(x_t(\xi_{[t]}), y_t(\xi_{[t]})) && \text{(4.6a)}\\
\text{s.t.} \quad & \big(x_1(\xi_{[1]}), x_2(\xi_{[2]}), \cdots, x_t(\xi_{[t]}), y_t(\xi_{[t]})\big) \in \mathcal{Z}_t \quad t = 1, \cdots, T, && \text{(4.6b)}\\
& r_{it_i^*} = 1 \Leftrightarrow \begin{cases} x_{it}(\xi_{[t]}) = \hat{x}_{it} \quad t = 1, \cdots, t_i^* - 1, \\ x_{it}(\xi_{[t]}) = \hat{x}_{it}(\xi_{[t_i^*]}) \quad t = t_i^*, \cdots, T, \end{cases} \qquad i = 1, \cdots, I, && \text{(4.6c)}\\
& \sum_{t=1}^{T} r_{it} = 1 \quad i = 1, \cdots, I, && \text{(4.6d)}\\
& r_{it} \in \{0, 1\} \quad t = 1, \cdots, T, i = 1, \cdots, I, && \text{(4.6e)}
\end{aligned}
$$

where the variable r_{it} is 1 if the state variable i has been revised at period t, and the function f_t corresponds to the objective function at period t. The set $\mathcal{Z}_t$ in (4.6b) represents the set of constraints corresponding to stage t. We consider a linear relationship for the constraints as described in the form (4.7), where the state decisions until period t is linked with the local decisions of that period. We note that these constraints can be also written in Markovian form by eliminating the state variables until period $t-1$:

$$\sum_{l=0}^{t-1} C_{t,t-l} x_{t-l}(\xi_{[t-l]}) + D_t y_t(\xi_{[t]}) \geq d_t. \tag{4.7}$$

We define the auxiliary decision variable $\hat{x}$ for constructing the adaptive relationship depending on the revision time of each state variable, as illustrated in constraint (4.6c). Thus, if $x_{it}(\xi_{[t]})$ is revised at stage t_i^*, then underlying data $\xi_{[t_i^*]}$ is observed at that period, and the decisions for the remainder of the planning horizon depend on those observations. Constraints (4.6d) and (4.6e) ensure that each state variable is revised once during the planning horizon.

As a next step, we focus on the case where data is random, in which we minimize the expected objective function value over the planning horizon. We denote the random data vector between stages t and t' as $\boldsymbol{\xi}_{[t,t']}$, and its realized value as $\xi_{[t,t']}$. For simplifying the notation, we denote the variable $x_t(\xi_{[t]})$ as x_t. In terms of the decision dynamics, we represent the adaptive two-stage stochastic case in a form where we determine the revision time of each state variable in the first level, and the second level turns into a multi-stage stochastic program once the revision times are fixed. Specifically, when the state variable x_{it} has been assigned to a revision time t_i^*, we determine $\{x_{it}\}_{t=1}^{t_i^*-1}$ at the beginning of the planning horizon. Next, we observe the underlying uncertainty, namely $\xi_{[t_i^*]}$, and determine the decisions x_{it} for the remainder of the planning horizon, accordingly. The resulting adaptive two-stage stochastic program can be represented in the following form:

$$\begin{aligned}\min_{r} \min_{(x_1,y_1)\in F_1} \{f_1(x_1,y_1) + \mathbb{E}_{\boldsymbol{\xi}_{[2,T]}|\xi_{[1]}}\big[\min_{(x_2,y_2)\in F_2(x_1,\xi_2,r)} \{f_2(x_2,y_2) + ... \\ + \mathbb{E}_{\boldsymbol{\xi}_{[T,T]}|\xi_{[T-1]}}\big[\min_{(x_T,y_T)\in F_T(x_{T-1},\xi_T,r)} \{f_T(x_T,y_T)\}\big]\}\big]\},\end{aligned}$$

where $F_t(x_{t-1},\xi_t,r)$ represents the feasible region of stage t, and $\mathbb{E}_{\boldsymbol{\xi}_{[t,T]}|\xi_{[t-1]}}$ is the expectation operator at stage t by realizing the observations until that period. As the revision decisions are fixed for the inner problem, we can represent the partially adaptive relationship in (4.6c) within the constraint set of each stage.

4.3.2 Scenario Tree Formulation

In order to represent the adaptive two-stage approach described in Section 4.3.1, we approximate the underlying stochastic process by generating finitely many samples. Specifically, scenario tree is a fundamental method to represent uncertainty in sequential decision making processes [74], where each node corresponds to a specific realization of the underlying uncertainty. To construct a scenario tree, sampling approaches similar to the ones proposed for Sample Average Approximation can be adopted to approximate the underlying uncertainty. Theoretical bounds and confidence intervals regarding the construction of these trees and optimality of the solutions are studied extensively in literature (see e.g. [75, 76, 77]).

In the remainder of this chapter, we consider a scenario tree $\mathcal{T}$ with T stages to model the uncertainty structure in a multi-period problem with T periods. We illustrate a sample scenario tree in Figure 4.2, and represent each node of the tree as $n \in \mathcal{T}$. We define the set of nodes in each period $1 \leq t \leq T$ as S_t, and the period of a node n as t_n. Each node n, except the root node, has an ancestor, which is denoted as $a(n)$. The unique path from the root node to a specific node n is represented by $P(n)$. We note that each path from the root node to a leaf node corresponds to a scenario, in other words each $P(n)$ gives a scenario when $n \in S_T$. We denote the subtree rooted at node n until period t as $\mathcal{T}(n,t)$ for $t_n \leq t \leq T$. To shorten the notation, when the last period of the subtree is T, we let $\mathcal{T}(n) := \mathcal{T}(n,T)$ for all $n \in \mathcal{T}$. The probability of each node n is given by p_n, where $\sum_{n \in S_t} p_n = 1$ for all $1 \leq t \leq T$.

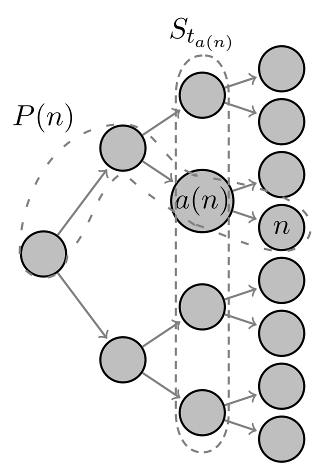


Figure 4.2: Scenario tree structure.

Utilizing the scenario tree structure introduced above, a general form multi-stage stochastic program can be formulated as

$$\min_{x,y} \quad \sum_{n\in\mathcal{T}} p_n(a_n^\top x_n + b_n^\top y_n) \tag{4.8a}$$

$$\text{s.t.} \quad \sum_{m\in P(n)} C_{nm}x_m + D_n y_n \geq d_n \quad \forall n \in \mathcal{T}, \tag{4.8b}$$

$$x_n \in \mathcal{X}_n, y_n \in \mathcal{Y}_n \quad \forall n \in \mathcal{T}, \tag{4.8c}$$

where the decision variables corresponding to node $n \in \mathcal{T}$ are given as x_n, y_n, and the parameters of this node is represented as (a_n, b_n, C_n, D_n, d_n). We denote the *state variables* as $\{x_n\}_{n\in\mathcal{T}}$, and the *stage variables* as $\{y_n\}_{n\in\mathcal{T}}$, where *stage variables* y_n are local variables to their associated stage t_n. Constraints referring to the variables x_n and y_n for each node $n \in \mathcal{T}$ are compactly represented by the sets $\mathcal{X}_n$ and $\mathcal{Y}_n$. We let the dimensions of the variables $\{x_n\}_{n\in\mathcal{T}}$ and $\{y_n\}_{n\in\mathcal{T}}$ be I and J, respectively. We note that constraint (4.8b) is analogous to constraint (4.7).

Two-stage stochastic programs are less adaptive compared to multi-stage approaches since they determine a single static solution for the stage variables per each period, irrespective of the specific realizations of that period. Consequently, two-stage stochastic programming version of the formulation (4.8) can be represented as

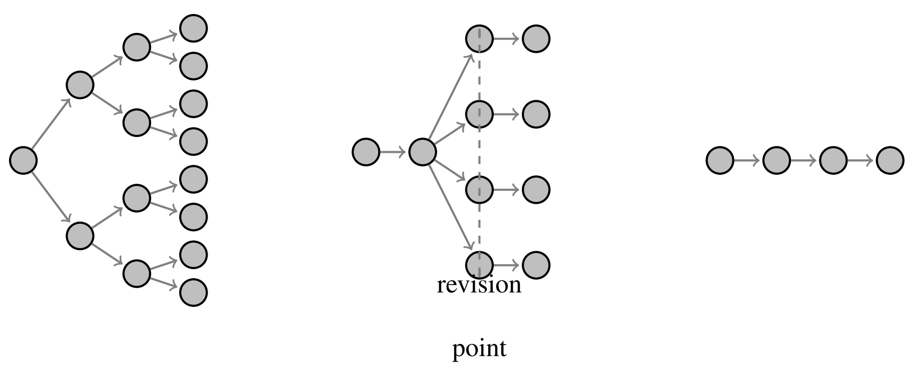
$$\min_{x,y} \quad \sum_{n\in\mathcal{T}} p_n(a_n^\top x_n + b_n^\top y_n), \tag{4.9a}$$

$$\text{s.t.} \quad (4.8\text{b}), (4.8\text{c})$$

$$x_m = x_n \quad \forall m, n \in S_t, t = 1, \cdots, T, \tag{4.9b}$$

where constraint (4.9b) ensures that the state variables $\{x_n\}_{n\in\mathcal{T}}$ are determined at the beginning of the planning horizon, and same across different scenarios. We allow $\{y_n\}_{n\in\mathcal{T}}$ decisions to have a multi-stage decision structure by allowing revisions with respect to the underlying uncertainty.

In this study, we focus on problem settings that require an intermediate approach between the multi-stage and two-stage stochastic programming by demanding a fixed sequence of decisions with a possible option to revise in future. Specifically, we propose the adaptive two-stage approach by optimizing the time to revise the decisions over a static policy. We illustrate this concept over scenario trees in Figure 4.3 by visualizing the decision structures of stochastic programming approaches for the state variable $\{x_n\}_{n \in \mathcal{T}}$. In multi-stage stochastic programming, decisions are specific to each node, whereas two-stage approaches provide a simpler method by resulting in one decision per each time period. We illustrate the adaptive two-stage decision structure by considering one revision point throughout the planning horizon. More specifically, we have a static decision structure before the revision stage. Once we revise our decisions, the remainder trees rooted from each node of the revision stage are compressed to have a single decision for the remaining planing horizon corresponding to that node. We note that the critical point of the adaptive two-stage approach is the choice of the revision point with respect to the underlying uncertainty.



(a) Multi-stage. (b) Adaptive two-stage. (c) Two-stage.

Figure 4.3: Decision structures for $\{x_n\}_{n \in \mathcal{T}}$ in different stochastic programming approaches.

We formalize the adaptive two-stage approach in (4.10) by allowing one revision decision for each state variable throughout the planing horizon. The revision times are denoted by the variable t^*.

$$\min_{x,y,t^*} \quad \sum_{n\in\mathcal{T}} p_n(\sum_{i=1}^{I} a_{in}x_{in} + \sum_{j=1}^{J} b_{jn}y_{jn}) \tag{4.10a}$$

$$\text{s.t.} \quad (4.8\text{b}), (4.8\text{c})$$

$$x_{im} = x_{in} \quad \forall m, n \in S_t, \quad t < t_i^*, i = 1, \cdots, I, \tag{4.10b}$$

$$x_{im} = x_{in} \quad \forall m, n \in S_t \cap \mathcal{T}(j), \quad j \in S_{t_i^*}, \quad t \geq t_i^*, i = 1, \cdots, I, \tag{4.10c}$$

$$t_i^* \in \{1, \cdots, T\} \quad \forall i \in I. \tag{4.10d}$$

Here, constraints (4.10b) and (4.10c) refer to the adaptive two-stage relationship under the revision decisions t_i^*, as illustrated in Figure 4.3 and similar to constraint (4.6c).

As constraints (4.10b) and (4.10c) depend on the decision variable t_i^*, we obtain a non-linear stochastic programming formulation in (4.10). To linearize this relationship, we introduce an auxiliary binary variable r_{it} for each $i \in I$, which is 1 if the decisions $\{x_{in}\}$ are revised at nodes $n \in S_t$, and 0 otherwise. Combining the above, we can reformulate the adaptive two-stage stochastic program as follows:

$$\min_{x,y,r} \quad \sum_{n\in\mathcal{T}} p_n(\sum_{i=1}^{I} a_{in}x_{in} + \sum_{j=1}^{J} b_{jn}y_{jn}) \tag{4.11a}$$

$$\text{s.t.} \quad (4.8\text{b}), (4.8\text{c})$$

$$\sum_{t=1}^{T} r_{it} = 1 \quad i = 1, \cdots, I, \tag{4.11b}$$

$$x_{im} \geq x_{in} - \bar{x}(1 - \sum_{t'=t+1}^{T} r_{it'}) \quad \forall m, n \in S_t, t = 1, \cdots, T-1, i = 1, \cdots, I, \tag{4.11c}$$

$$x_{im} \leq x_{in} + \bar{x}(1 - \sum_{t'=t+1}^{T} r_{it'}) \quad \forall m, n \in S_t, t = 1, \cdots, T-1, i = 1, \cdots, I, \tag{4.11d}$$

$$x_{im} \geq x_{in} - \bar{x}(1 - r_{it}) \quad \forall m, n \in S_{t'} \cap \mathcal{T}(l), l \in S_t, t' \geq t, t = 1, \cdots, T, i = 1, \cdots, I, \tag{4.11e}$$

$$x_{im} \leq x_{in} + \bar{x}(1 - r_{it}) \quad \forall m, n \in S_{t'} \cap \mathcal{T}(l), l \in S_t, t' \geq t, t = 1, \cdots, T, i = 1, \cdots, I, \tag{4.11f}$$

$$r_{it} \in \{0, 1\} \quad \forall i = 1, \cdots, I, t = 1, \cdots, T. \tag{4.11g}$$

Here, constraint (4.11b) ensures that each state variable is revised once throughout the planning horizon. Constraints (4.11c) and (4.11d) guarentee that the decision of each state variable until its revision point is determined at the beginning of the planning horizon. Constraints (4.11e) and (4.11f) correspond to the decisions after the revision point by observing the underlying uncertainty at that time. We note that the parameter $\bar{x}$ denotes an upper bound value on the decision variables $\{x_{in}\}_{n \in \mathcal{T}}$ for each $i = 1, \cdots, I$, whose specific value depends on the problem structure.

Although optimization model in (4.11) provides a stochastic mixed-integer linear programming formulation for the adaptive two-stage approach, it requires the addition of exponentially many linear constraints in terms of the number of periods for representing the desired relationship. Additionally, constraints (4.11c)–(4.11f) involve big-M coefficients which may weaken the corresponding linear programming relaxation. Consequently, it is computationally challenging to directly solve the formulation (4.11) when the size of the tree becomes larger (see, for instance, Tables 4.2 and 4.3 in Section 4.6). This computational complexity associated with the adaptive two-stage formulation has motivated us to understand the theoretical and empirical performances of the proposed methodology on specific problem structures.

4.3.3 Complexity

To understand the difficulty of this problem, we identify its computational complexity.

Theorem 3. *Solving the adaptive two-stage stochastic programming model in* (4.10) *is NP-Hard.*

Proof. To prove the theorem, it suffices to show that the feasibility problem associated with the adaptive two-stage problem (4.10) is $\mathcal{NP}$-Complete. For this purpose, we define the feasible region constructed in (4.12), and refer to this feasibility problem as P. Next, we demonstrate that the subset sum problem, which is known to be $\mathcal{NP}$-Complete [78], can be reduced to P in polynomial time. We specify the subset sum problem as follows: Given the non-negative integers $w'_1, w'_2, \cdots, w'_{N'}$, and W', does there exist a subset $S \subseteq \{1, \cdots, N'\}$ such that $\sum_{i \in S} w'_i = W'$?

$$\min_{\alpha,\beta,t^*} \quad 0 \tag{4.12a}$$

$$\text{s.t.} \quad \alpha_{i2} - \beta_{i2} = \alpha_{i1} \qquad i = 1, \cdots, N \tag{4.12b}$$

$$\alpha_{i3} - \beta_{i3} = -\alpha_{i1} \qquad i = 1, \cdots, N \tag{4.12c}$$

$$\alpha_{i4} - \beta_{i4} = \alpha_{i1} \qquad i = 1, \cdots, N \tag{4.12d}$$

$$\alpha_{i5} - \beta_{i5} = 2 - \alpha_{i1} \qquad i = 1, \cdots, N \tag{4.12e}$$

$$\alpha_{i6} - \beta_{i6} = \alpha_{i1} \qquad i = 1, \cdots, N \tag{4.12f}$$

$$\alpha_{i7} - \beta_{i7} = 2 - \alpha_{i1} \qquad i = 1, \cdots, N \tag{4.12g}$$

$$\sum_{i=1}^{N} w_i \alpha_{i1} = W \tag{4.12h}$$

$$\alpha_{im} = \alpha_{in}, \beta_{im} = \beta_{in} \qquad \forall m, n \in S_t, \quad t < t_i^*, i = 1, \cdots, N \tag{4.12i}$$

$$\alpha_{im} = \alpha_{in}, \beta_{im} = \beta_{in} \quad \forall m, n \in S_t \cap \mathcal{T}(j), \quad j \in S_{t_i^*}, \quad t \geq t_i^*, i = 1, \cdots, N \tag{4.12j}$$

$$t_i^* \in \{1, 2, 3\} \qquad i = 1, \cdots, N. \tag{4.12k}$$

Clearly, the problem (4.12) is an instance of the adaptive two-stage problem. Specifically, we let $I = N$, $J = 0$, and $x_{in} = \begin{bmatrix} \alpha_{in} \\ \beta_{in} \end{bmatrix}$. We consider the scenario tree $\mathcal{T}$ with $T = 3$ stages as depicted in Figure 4.4. We let $\mathcal{T} = \{1, \cdots, 7\}$ be the set of nodes in ascending order, where node 1 represents the root node. We also let $N = N'$, $W = W'$, and $w_i = w_i'$ for all $i = 1, \cdots, N$. We note that constraints (4.12b) - (4.12g) correspond to constraint (4.8b), and constraint (4.12h) represent the constraint (4.8c). Similarly, constraints (4.12i), (4.12j) and (4.12k) refer to the constraints (4.10b), (4.10c) and (4.10d), respectively.

Lemma 2. *The following hold for the problem* (4.12)*:*

1. *If $t_i^* = 1$ for any $i = 1, \cdots, N$, then the problem* (4.12) *is infeasible.*
2. *If $t_i^* = 2$, then $\alpha_{i1} = 1$ for all $i = 1, \cdots, N$.*
3. *If $t_i^* = 3$, then $\alpha_{i1} = 0$ for all $i = 1, \cdots, N$.*

The proof of the Lemma 2 follows from the construction of the program (4.12). Specifically, when $t_i^* = 1$ for any $i = 1, \cdots, N$, then due to constraints (4.12i) and (4.12j), $\alpha_{i1} = -\alpha_{i1} = 2 - \alpha_{i1}$ resulting in infeasibility. When $t_i^* = 2$, then we have $\alpha_{i1} = 2 - \alpha_{i1}$ due to constraints (4.12d) and (4.12g), resulting in $\alpha_{i1} = 1$. Similarly, when $t_i^* = 3$, we have $\alpha_{i1} = \alpha_{i1}$ due to constraints (4.12b) and (4.12c) making $\alpha_{i0} = 0$.

We first show that given a solution S to the subset-sum problem, we can construct a feasible solution for problem P. Specifically, we set $\alpha_{i1} = 1$ if $i \in S$, and 0 otherwise

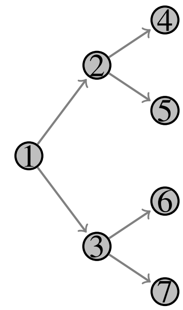


Figure 4.4: Scenario tree for the problem (4.12).

for all $i = 1, \cdots, N$. If $\alpha_{i1} = 1$, then $t_i^* = 2$ resulting in $\alpha_{i2} = 1$, $\beta_{i2} = 0$, $\alpha_{i3} = 0$, $\beta_{i3} = 1$, $\beta_{i4} = \beta_{i5} = \beta_{i6} = \beta_{i7} = 0$, $\alpha_{i4} = \alpha_{i5} = \alpha_{i6} = \alpha_{i7} = 1$ for all $i = 1, \cdots, N$. If $\alpha_{i1} = 0$, then $t_i^* = 3$ resulting in $\alpha_{i2} = \beta_{i2} = \alpha_{i3} = \beta_{i3} = 0$, $\alpha_{i4} = \beta_{i4} = 0$, $\alpha_{i5} = 2$, $\beta_{i5} = 0$, $\alpha_{i6} = 0$, $\beta_{i6} = 0$, $\alpha_{i7} = 2$, $\beta_{i7} = 0$. Secondly, given a feasible solution (α, β, t^*) to the problem P, we can construct a feasible solution S for the subset-sum. We note that in a feasible solution of P, we have either t_i^* = 2 or $t_i^* = 3$ for $i = 1, \cdots, N$, as noted in Lemma 2. To construct a feasible solution, we select $S = \{i : t_i^* = 2\}$. As (α, β, t^*) is a feasible solution of P, we satisfy the condition $\sum_{i \in S} w'_i = W'$. Combining the above, we prove that P is $\mathcal{NP}$-Complete, completing the proof. □

The proof of Theorem 3 demonstrates that the hardness of the adaptive two-stage problem comes from the choice of the revision times, i.e. t_i^* for all $i \in I$. In particular, the adaptive two-stage problem considered in the proof reduces to a linear program with polynomial size when revision time decisions are fixed, becoming polynomially solvable.

4.3.4 Value of adaptive two-stage stochastic solutions

In this section, we present a way to assess the performances of the stochastic programming approaches according to their adaptiveness level to uncertainty under a given revision decision $t^* \in \mathbb{Z}_+^I$. Specifically, we have the following relationship for the vector $t^* \in \{1, \cdots, T\}^I$

$$V^{MS} \leq V^{ATS}(t^*) \leq V^{TS},$$

where V^{MS}, $V^{ATS}(t^*)$, and V^{TS} correspond to the objective values of the multi-stage, adaptive two-stage under given revision decision t^*, and two-stage models, respectively. We note that this relation holds as the two-stage program provides a feasible solution for the adaptive two-stage program under any t^* vector, and the solution of the adaptive two-stage program under any t^* vector is feasible for the multi-stage program.

In order to evaluate the performance of the adaptive two-stage approach, we aim deriving bounds for the $V^{MS} - V^{ATS}(t^*)$, and $V^{TS} - V^{ATS}(t^*)$ for a given t^* vector. We refer to the bound $V^{TS} - V^{ATS}(t^*)$ as *value of adaptive two-stage* (VATS) in the remainder of this chapter. We note that in Section 4.2.2, we illustrate the value of the adaptive two-stage approach in comparison to static and fully dynamic policies over a multi-period newsvendor problem. Here, static and fully dynamic policies align with two-stage and multi-stage stochastic programming approaches, respectively. Our analysis demonstrates that the optimal objective value is significantly affected by the choice of the revision time, making it critical to determine the best time to realize the underlying uncertainty.

4.4 Analysis of Capacity Expansion Planning Problem

In this section, we study a special problem structure where analytical bounds can be derived to compare the solution performances of two-stage, multi-stage and adaptive two-stage approaches. The structure studied is relevant for a class of problems including capacity expansion and investment planning. In the remainder of this section, we first present the problem formulation for the adaptive two-stage approach under given revision points. Then, we provide a theoretical analysis of the performance of the proposed methodology with respect to the choice of the revision decisions.

4.4.1 Problem formulation

The capacity expansion problem determines the capacity acquisition decisions of the set of resources $\mathcal{I}$ by allocating the corresponding capacities to the tasks $\mathcal{J}$, while satisfying the demand of items $\mathcal{K}$ and capacity constraints. A stochastic capacity expansion problem with T periods can be written as follows:

$$\min_{x,y} \quad \sum_{n\in\mathcal{T}} p_n(a_n^\top x_n + b_n^\top y_n) \tag{4.13a}$$

$$\text{s.t.} \quad A_n y_n \leq \sum_{m\in P(n)} x_m \quad \forall n \in \mathcal{T}, \tag{4.13b}$$

$$B_n y_n \geq d_n \quad \forall n \in \mathcal{T}, \tag{4.13c}$$

$$x_n \in \mathbb{Z}_+^{|\mathcal{I}|}, y_n \in \mathbb{R}_+^{|\mathcal{J}|} \quad \forall n \in \mathcal{T}, \tag{4.13d}$$

where the decision variables x_n, y_n, and the parameter p_n represent the capacity acquisition decisions, capacity allocation decisions, and probability corresponding to the node n, respectively. By adopting the scenario tree structure, we consider the problem parameters as $(a_n, b_n, A_n, B_n, d_n)$ corresponding to the node n. Objective (4.13a) minimizes the total cost by considering capacity acquisition and allocation decisions. Constraint (4.13b) guarantees that the assigned capacity in each period is less than the available capacity, and constraint (4.13c) ensures that the demand is satisfied.

We note that the proposed multi-stage stochastic problem (4.13) can be reformulated as in [79] by considering its specific substructure. The resulting formulation can be stated as follows:

$$\min_{y} \quad \sum_{n\in\mathcal{T}} p_n b_n^\top y_n + \sum_{i\in\mathcal{I}} Q_i(y) \tag{4.14a}$$

$$\text{s.t.} \quad (4.13c)$$

$$y_n \in \mathbb{R}_+^{|\mathcal{J}|} \quad \forall n \in \mathcal{T}, \tag{4.14b}$$

where, for each $i \in \mathcal{I}$,

$$Q_i(y) = \min_{x_i} \quad \sum_{n\in\mathcal{T}} p_n a_{in} x_{in}, \tag{4.15a}$$

$$\text{s.t.} \quad \sum_{m\in P(n)} x_{im} \geq [A_n y_n]_i \quad \forall n \in \mathcal{T}, \tag{4.15b}$$

$$x_{in} \in \mathbb{Z}_+ \quad \forall n \in \mathcal{T}. \tag{4.15c}$$

In the proposed reformulation, the first problem (4.14) determines the capacity allocation decisions, whereas the second problem (4.15) considers the capacity acquisition decisions given the allocation decisions, which is further decomposed with respect to each item $i \in \mathcal{I}$. We let $I = |\mathcal{I}|$ in the remainder of this section.

We note that under a given sequence of allocation decisions $\{y_n\}_{n\in\mathcal{T}}$, we can obtain the optimal capacity acquisition decisions corresponding to each item i by solving (4.15). This problem can be restated by letting $\delta_{in} = [A_n y_n]_i$ as follows:

$$\min_{x_i} \quad \sum_{n\in\mathcal{T}} p_n a_{in} x_{in} \tag{4.16a}$$

$$\text{s.t.} \quad \sum_{m\in P(n)} x_{im} \geq \delta_{in} \quad \forall n \in \mathcal{T} \tag{4.16b}$$

$$x_{in} \in \mathbb{Z}_+ \quad \forall n \in \mathcal{T}. \tag{4.16c}$$

We first observe that the feasible region of the formulation (4.16) gives an integral polyhedron under integer $\{\delta_{in}\}_{n\in\mathcal{T}}$ values [79]. Furthermore, it is equivalent to a simpler version of the stochastic lot sizing problem without a fixed order cost.

The two-stage version of the single-resource problem (4.16) can be represented as follows:

$$
\begin{aligned}
\min_{x_i} \quad & \sum_{n \in \mathcal{T}} p_n a_{in} x_{in} && (4.17a)\\
\text{s.t.} \quad & (4.16b), (4.16c) \\
& x_{im} = x_{in} \quad \forall m, n \in S_t. && (4.17b)
\end{aligned}
$$

As an alternative to the formulations (4.16) and (4.17), we consider the adaptive two-stage stochastic programming approach where the capacity allocation decision of each item $i \in \mathcal{I}$ is determined at the beginning of planning until stage $t_i^* - 1$. Then, the underlying scenario tree at time t_i^* is observed, and the decisions until the end of the planning horizon are determined at that time. The resulting problem can be formulated as follows:

$$
\begin{aligned}
\min_{x_i} \quad & \sum_{n \in \mathcal{T}} p_n a_{in} x_{in} && (4.18a)\\
\text{s.t.} \quad & (4.16b), (4.16c) \\
& x_{im} = x_{in} \quad \forall m, n \in S_t, \quad t < t_i^* && (4.18b)\\
& x_{im} = x_{in} \quad \forall m, n \in S_t \cap \mathcal{T}(j), \quad j \in S_{t_i^*}, \quad t \geq t_i^*. && (4.18c)
\end{aligned}
$$

We note that the above formulation considers the case under a given t_i^* value. We can equivalently represent (4.18) by defining a condensed scenario tree based on the t_i^* value. Let the condensed version of tree $\mathcal{T}$ for item i be $\mathcal{T}_i(t_i^*)$. We denote the set of nodes that are condensed to node $n \in \mathcal{T}_i(t_i^*)$ by $\hat{C}_n \subset \mathcal{T}$, where $\hat{C}_n$ is obtained by combining the nodes with the same decision structure as in constraints (4.18b) and (4.18c). Then, the resulting formulation is

$$\begin{aligned} \min_{x_i} \quad & \sum_{n \in \mathcal{T}_i(t_i^*)} \hat{p}_n \hat{a}_{in} x_{in} && \text{(4.19a)} \\ \text{s.t.} \quad & \sum_{m \in P(n)} x_{im} \geq \hat{\delta}_{in} \quad \forall n \in \mathcal{T}_i(t_i^*) && \text{(4.19b)} \\ & x_{in} \in \mathbb{Z}_+ \quad \forall n \in \mathcal{T}_i(t_i^*), && \text{(4.19c)} \end{aligned}$$

where $\hat{p}_n = \sum_{m \in \hat{C}_n} p_m$, $\hat{a}_{in} = \sum_{m \in \hat{C}_n} p_m a_{im}$, and $\hat{\delta}_{in} = \max_{m \in \hat{C}_n} \{\delta_{im}\}$.

Proposition 6. *The coefficient matrix of the formulation* (4.19) *is totally unimodular.*

Proof. We observe that the coefficient matrix of (4.19) is composed of 0 and 1 values. Additionally, number of 1's in each column j is equal to $|\mathcal{T}_i(t_i^*)(j)|$, where $\mathcal{T}_i(t_i^*)(j)$ represents the subtree rooted at node j. More specifically, an entry corresponding to i^{th} row and j^{th} column of the coefficient matrix is 1 if and only if node $i \in \mathcal{T}_i(t_i^*)(j)$. Consequently, we can rearrange the coefficient matrix to obtain an interval matrix. This demonstrates that the desired matrix is totally unimodular. □

4.4.2 Deriving VATS for Capacity Expansion Planning

To evaluate the performance of the adaptive two-stage approach, we aim deriving bounds for the $V^{MS} - V^{ATS}(t^*)$, and $V^{TS} - V^{ATS}(t^*)$ for the capacity expansion problem (4.13) under a given revision vector t^*. We first study the single item version of the subproblem (4.15) under different stochastic programming formulations in Section 4.4.2, and present a sensitivity analysis on these bounds in Section 4.4.2. Then, we extend these results to the complete capacity expansion problem in Section 4.4.2.

VATS for the Single-Resource Problem

In this section, we derive the VATS for the single-resource problem (4.16) using its linear programming relaxations under two-stage, adaptive two-stage and multi-stage models. We let v^M, v^T, $v^R(t^*)$ be the optimal value of the linear programming relaxations of the formulations (4.16), (4.17), and (4.18) respectively. As we focus on the single-resource problem, we omit the resource index i in this section for brevity.

To construct the VATS, we examine the problem parameters with respect to the underlying scenario tree. Specifically, we define the minimum and maximum costs over the scenario tree as $a_* = \min_{n\in\mathcal{T}}\{a_n\}$ and $a^* = \max_{n\in\mathcal{T}}\{a_n\}$, respectively. We denote the maximum demand over the scenario tree as $\delta^* = \max_{n\in\mathcal{T}}\{\delta_n\}$, and the expected maximum demand over scenarios as $\bar{\delta} = \sum_{n\in S_T} p_n \max_{m\in P(n)}\{\delta_m\}$. Additionally, we examine the problem parameters based on the choice of the revision time. In particular, we define the minimum and maximum cost parameters before and after the revision time t^* as follows:

$$\underline{a}^-(t^*) = \min_{n\in\mathcal{T}:t_n<t^*}\{a_n\}, \quad \underline{a}^+(t^*) = \min_{n\in\mathcal{T}:t_n\geq t^*}\{a_n\},$$
$$\bar{a}^-(t^*) = \max_{n\in\mathcal{T}:t_n<t^*}\{a_n\}, \quad \bar{a}^+(t^*) = \max_{n\in\mathcal{T}:t_n\geq t^*}\{a_n\}.$$

For the demand parameters, we let

$$\delta^-(t^*) = \max_{m\in\mathcal{T}(1,t^*-1)}\{\delta_m\}, \quad \delta^+(t^*) = \sum_{n\in S_{t^*}} p_n \max_{m\in\mathcal{T}(1,t^*-1)\cup\mathcal{T}(n)}\{\delta_m\}.$$

Here, the parameter $\delta^-(t^*)$ represents the maximum demand value before the revision time t^*. The parameter $\delta^+(t^*)$ corresponds to the expected maximum demand over the tree until the revision time t^* and the subtree rooted at each node of the revision stage. By comprehending the underlying uncertainty through these definitions, we obtain our main result for analyzing the single-resource problem.

Theorem 4. *We derive the following bounds for $v^T - v^R(t^*)$ and $v^R(t^*) - v^M$:*

$$\begin{aligned} a_*\delta^* - (\bar{a}^-(t^*) - \bar{a}^+(t^*))\delta^-(t^*) - \bar{a}^+(t^*)\delta^+(t^*) &\leq v^T - v^R(t^*) \\ &\leq a^*\delta^* - (\underline{a}^-(t^*) - \underline{a}^+(t^*))\delta^-(t^*) - \underline{a}^+(t^*)\delta^+(t^*), \end{aligned} \tag{4.20}$$

and

$$(\underline{a}^-(t^*) - \underline{a}^+(t^*))\delta^-(t^*) + \underline{a}^+(t^*)\delta^+(t^*) - a^*\bar{\delta} \le v^R(t^*) - v^M \qquad (4.21)$$
$$\le (\bar{a}^-(t^*) - \bar{a}^+(t^*))\delta^-(t^*) + \bar{a}^+(t^*)\delta^+(t^*) - a_*\bar{\delta}.$$

To obtain the value of the adaptive two-stage approach in Theorem 4, we first identify the bounds on the optimum values v^M, v^T, $v^R(t^*)$.

Proposition 7. *We can derive the following bound for $v^R(t^*)$:*

$$(\underline{a}^-(t^*)-\underline{a}^+(t^*))\delta^-(t^*)+\underline{a}^+(t^*)\delta^+(t^*) \le v^R(t^*) \le (\bar{a}^-(t^*)-\bar{a}^+(t^*))\delta^-(t^*)+\bar{a}^+(t^*)\delta^+(t^*).$$

Proof. The decision variables $\{x_n\}_{\{n\in\mathcal{T}\}}$ can be categorized into two groups according to the revision time t^*. First of all, for $t < t^*$, the solutions $\{x_m : m \in S_t\}$ have the same value. Thus, we denote variables until period t^* as $\hat{x}_t$ for all $t = 1, \cdots, t^* - 1$. Secondly, for $t \ge t^*$, we have the same solutions for $\{x_m : m \in S_t \cap \mathcal{T}(n)\}$ for all $n \in S_{t^*}$. We refer to these variables as $\hat{x}_{nt}$ for any $n \in S_{t^*}$ and $t \ge t^*$.

In order to find an upper bound on $v^R(t^*)$, we construct a feasible solution where $\hat{x}_1 = \delta_1$, $\hat{x}_t = \max_{m\in\mathcal{T}(1,t)}\{\delta_m\} - \max_{m\in\mathcal{T}(1,t-1)}\{\delta_m\}$ for $2 \le t \le t^* - 1$, and $\hat{x}_{nt} = \max\{\delta_m : m \in \mathcal{T}(1, t^* - 1) \cup \mathcal{T}(n, t)\} - \max\{\delta_m : m \in \mathcal{T}(1, t^* - 1) \cup \mathcal{T}(n, t - 1)\}$ for $t \ge t^*$. Using the relationship $v^R(t^*) \le \sum_{n\in\mathcal{T}} p_n a_n x_n$ for any feasible $\{x_n\}_{\{n\in\mathcal{T}\}}$, we can obtain the following:

$$\begin{aligned}
v^R(t^*) &\le \sum_{t=1}^{t^*-1}\sum_{n\in S_t} p_n a_n \hat{x}_t + \sum_{t=t^*}^{T}\sum_{n\in S_{t^*}} p_n a_n \hat{x}_{nt} \\
&\le \bar{a}^-(t^*)\sum_{t=1}^{t^*-1}\hat{x}_t + \bar{a}^+(t^*)\sum_{n\in S_{t^*}} p_n \sum_{t=t^*}^{T}\hat{x}_{nt} \\
&= \bar{a}^-(t^*)\sum_{t=1}^{t^*-1}\big(\max_{m\in\mathcal{T}(1,t)}\{\delta_m\} - \max_{m\in\mathcal{T}(1,t-1)}\{\delta_m\}\big)
\end{aligned}$$

$$
\begin{aligned}
&+ \bar{a}^+(t^*) \sum_{n \in S_{t^*}} p_n \sum_{t=t^*}^{T} (\max_{m \in \mathcal{T}(1,t^*-1) \cup \mathcal{T}(n,t)} \{\delta_m\} - \max_{m \in \mathcal{T}(1,t^*-1) \cup \mathcal{T}(n,t-1)} \{\delta_m\}) \\
&= \bar{a}^-(t^*) \max_{m \in \mathcal{T}(1,t^*-1)} \{\delta_m\} + \bar{a}^+(t^*) \sum_{n \in S_{t^*}} p_n (\max_{m \in \mathcal{T}(1,t^*-1) \cup \mathcal{T}(n)} \{\delta_m\} - \max_{m \in \mathcal{T}(1,t^*-1)} \{\delta_m\}) \\
&= \bar{a}^-(t^*) \max_{m \in \mathcal{T}(1,t^*-1)} \{\delta_m\} + \bar{a}^+(t^*) \sum_{n \in S_{t^*}} p_n \max_{m \in \mathcal{T}(1,t^*-1) \cup \mathcal{T}(n)} \{\delta_m\} - \bar{a}^+(t^*) \max_{m \in \mathcal{T}(1,t^*-1)} \{\delta_m\} \\
&= (\bar{a}^-(t^*) - \bar{a}^+(t^*))\delta^-(t^*) + \bar{a}^+(t^*)\delta^+(t^*).
\end{aligned}
$$

Let $\hat{x}_t^*$ for $t < t^*$ and $\hat{x}_{nt}^*$, for $t \geq t^*$ and $n \in S_{t^*}$ denote the optimal solution for the problem (4.18). Then, we can derive the following:

$$
\begin{aligned}
v^R(t^*) &= \sum_{t=1}^{t^*-1} \sum_{n \in S_t} p_n a_n \hat{x}_t^* + \sum_{t=t^*}^{T} \sum_{n \in S_{t^*}} p_n a_n \hat{x}_{nt}^* \\
&\geq \underline{\mathrm{a}}^-(t^*) \sum_{t=1}^{t^*-1} \hat{x}_t^* + \underline{\mathrm{a}}^+(t^*) \sum_{n \in S_{t^*}} p_n \sum_{t=t^*}^{T} \hat{x}_{nt}^* \\
&\geq \underline{\mathrm{a}}^-(t^*) \sum_{t=1}^{t^*-1} \hat{x}_t^* + \underline{\mathrm{a}}^+(t^*) \sum_{n \in S_{t^*}} p_n (\max_{m \in \mathcal{T}(1,t^*-1) \cup \mathcal{T}(n)} \{\delta_m\} - \sum_{t=1}^{t^*-1} \hat{x}_t^*) \\
&= (\underline{\mathrm{a}}^-(t^*) - \underline{\mathrm{a}}^+(t^*)) \sum_{t=1}^{t^*-1} \hat{x}_t^* + \underline{\mathrm{a}}^+(t^*) \sum_{n \in S_{t^*}} p_n (\max_{m \in \mathcal{T}(1,t^*-1) \cup \mathcal{T}(n)} \{\delta_m\}) \\
&\geq (\underline{\mathrm{a}}^-(t^*) - \underline{\mathrm{a}}^+(t^*))\delta^-(t^*) + \underline{\mathrm{a}}^+(t^*)\delta^+(t^*).
\end{aligned}
$$

The second and fourth inequalities follow from constraint (4.16b). In particular, any feasible solution $\hat{x}_t$ and $\hat{x}_{nt}$ for any $n \in S_{t^*}$ should satisfy $\sum_{t=1}^{t^*-1} \hat{x}_t + \sum_{t=t^*}^{T} \hat{x}_{nt} \geq \max_{m \in \mathcal{T}(1,t^*-1) \cup \mathcal{T}(n)} \{\delta_m\}$. Additionally, $\sum_{t=1}^{t^*-1} \hat{x}_t \geq \max_{m \in \mathcal{T}(1,t^*-1)} \{\delta_m\} = \delta^-(t^*)$.

Combining above, we derive upper and lower bounds to $v^R(t^*)$. □

We note that we can obtain bounds for the linear programming relaxation of the two-stage equivalent of the formulation (4.16) using Proposition 7 by selecting the revision period t^* as 1.

Corollary 1. *For v^T, we have $a_*\delta^* \leq v^T \leq a^*\delta^*$.*

Combining Proposition 7 and Corollary 1, we can obtain the VATS for the single-resource problem in (4.20) as in the first part of Theorem 4.

Proposition 8. *[79] For v^M, we have $a_*\bar{\delta} \leq v^M \leq a^*\bar{\delta}$.*

Combining Propositions 7 and 8, we obtain the value of the multi-stage model against adaptive two-stage model for the single-resource problem in (4.21) as in the second part of Theorem 4.

Sensitivity Analysis on the Single-Resource Problem

In order to gain some insight regarding the effect of the revision point decisions on the performance of the analytical bounds, we consider two cases with respect to the values of cost and demand parameters.

Demand Sensitivity We first consider the case where the cost parameters $\{a_n\}_{n\in\mathcal{T}}$ are almost equal to each other across the scenario tree, i.e. $\{a_n\}_{n\in\mathcal{T}} \approx a$. In that case, the bounds (4.20) and (4.21) reduce to the following expressions:

$$v^T - v^R(t^*) \approx a(\delta^* - \delta^+(t^*)), \quad v^R(t^*) - v^M \approx a(\delta^+(t^*) - \bar{\delta}). \tag{4.22}$$

Consequently, the value of the adaptive formulation depends on how much $\delta^+(t^*)$ differs from the maximum demand value, δ^*. Similarly, the value of the multi-stage stochastic program against the adaptive two-stage approach depends on the difference between $\delta^+(t^*)$ and the maximum average demand, $\bar{\delta}$. These results highlight the importance of the variability of the demand on the values of the stochastic programs. If there is not much variability in the demand across scenarios, then the three approaches result in similar solutions as the bounds in (4.22) go to zero. As variability of the scenarios increases, the corresponding bounds might change accordingly.

Furthermore, the bounds (4.22) demonstrate that the values of the adaptive approach might be highly dependent to the selection of t^* value. In particular, best performance gains for the adaptive approach, compared to two-stage and multi-stage cases, can be obtained when $\delta^+(t^*)$ is minimized. By this way, we obtain the solution of the adaptive program which gives the least loss compared to the multi-stage stochastic programs, and the most gain compared to the two-stage stochastic programs, with respect to the proposed bounds. Let us denote the best revision point in terms of the demand bounds as $t^{DB} := \operatorname{argmin}_{2\le t\le T} \delta^+(t)$.

We illustrate the effect of the revision times on the bound values (4.22) for an instance of the single-resource problem described in Appendix A.1. Specifically, we have unit cost for $\{a_n\}_{n\in\mathcal{T}}$ values, and $\{\delta_n\}_{n\in\mathcal{T}}$ values are sampled from a probability distribution. We present the bound values in Table 4.1, where $\delta^* = 41$ and $\bar{\delta} = 34.06$, considering the scenario tree in Figure A.1. We observe that $\delta^+(t^*)$ is minimized at period 3, i.e. $t^{DB} = 3$, maximizing the objective gap of the adaptive two-stage approach with two-stage model and minimizing the corresponding gap with multi-stage model. These results demonstrate that we can analytically determine the revision points for the single-resource problem when costs remain same during the planning horizon.

Table 4.1: Bound comparison with respect to the revision time.

	$t^* = 1$	$t^* = 2$	$t^* = 3$	$t^* = 4$	$t^* = 5$
$v^T - v^R(t^*)$	0.00	1.50	**5.25**	2.63	0.00
$v^R(t^*) - v^M$	6.94	5.44	**1.69**	4.31	6.94
$\delta^+(t^*)$	41.00	39.50	**35.75**	38.38	41.00

Next, we examine how $\delta^+(t^*)$ changes under different δ_n structures. Let $\hat{t}^D$ be the first stage that we observe the maximum demand value over the scenario tree, i.e. $\hat{t}^D = \min\{t \in \{1, \cdots, T\} : \delta_n = \delta^*, n \in S_t\}$.

Proposition 9. *Under a general demand structure for δ_n with cost values $\{a_n\}_{n\in\mathcal{T}} \approx a$, if*

$\hat{t}^D = 1$, *then* $t^{DB} \in \{2, \cdots, T\}$. *Otherwise,* $t^{DB} \in \{2, \cdots, \hat{t}^D\}$.

Proof. If $\hat{t}^D = 1$, then $\delta^+(t) = \delta^*$ for any $t \in \{1, \cdots, T\}$. Otherwise, $\delta^+(1) = \delta^*$ as $\mathcal{T}(1)$ corresponds to the full scenario tree including the maximum demand value. For any node n in stages $t > \hat{t}^D$, we have $\max_{m \in \mathcal{T}(1,t-1) \cup \mathcal{T}(n)}\{\delta_m\} = \delta^*$. Hence, we obtain $\delta^+(t) = \delta^*$ for $t > \hat{t}^D$. For an intermediate stage t between 1 and $\hat{t}^D$, we have the relationship $\delta^+(t) \leq \delta^*$. Combining the above, under a general demand structure for δ_n when $\hat{t}^D > 1$, we conclude that the minimizer of $\delta^+(t)$ is in $\{2, \cdots, \hat{t}^D\}$. $\square$

For specific forms of demand patterns, we can further refine Proposition 9.

Proposition 10. *Let stage* $t_n \in \{2, \cdots, T\}$ *has* N_t *many independent realizations for the demand values* $\{\delta_n\}_{n \in \mathcal{T}}$ *where cost values* $\{a_n\}_{n \in \mathcal{T}} \approx a$. *Then,* $t^{DB} = \hat{t}^D$.

Proof. We observe that for the cases $t < \hat{t}^D$ and $t > \hat{t}^D$, we have $\delta^+(t) = \delta^*$. For $t = t^D$, we obtain the relationship $\max_{m \in \mathcal{T}(1,t-1) \cup \mathcal{T}(n)}\{\delta_m\} < \delta^*$ for some $n \in S_{\hat{t}^D}$. Hence, we have $\delta^+(\hat{t}^D) < \delta^*$ making $\hat{t}^D$ the minimizer of $\delta^+(t)$. $\square$

Cost Sensitivity Next, we consider the case where the demand parameters $\{\delta_n\}_{n \in \mathcal{T}}$ are almost equal to each other throughout the scenario tree, i.e. $\{\delta_n\}_{n \in \mathcal{T}} \approx \delta$. In that case, the bounds (4.20) and (4.21) reduce to the following

$$\max\{(a_* - \bar{a}^-(t^*))\delta, 0\} \lessapprox v^T - v^R(t^*) \lessapprox (a^* - \underline{a}^-(t^*))\delta, \tag{4.23}$$

$$\max\{(\underline{a}^-(t^*) - a^*)\delta, 0\} \lessapprox v^R(t^*) - v^M \lessapprox (\bar{a}^-(t^*) - a_*)\delta. \tag{4.24}$$

We observe that the value of the adaptive formulation depends on the choice of the revision point as the bounds in (4.23) and (4.24) are functions of t^*. In order to select the best revision point for the adaptive two-stage approach, we aim maximizing the lower bound on its objective difference between two-stage in (4.23) and minimize the upper bound on the corresponding difference between multi-stage in (4.24). Thus, this results in finding the revision point $2 \leq t^* \leq T$ that minimizes $\bar{a}^-(t^*) - a_*$. Since a_* is not dependent to the choice of the revision point, it is omitted in the remainder of our analysis. We denote the best revision point in terms of the cost bounds as $t^{CB} := \operatorname{argmin}_{2\leq t\leq T} \bar{a}^-(t)$.

We consider several specific cases to illustrate the revision point decision t^{CB} based on the cost values. If cost values are increasing in each stage of the scenario tree, then $\bar{a}^-(t)$ is monotonically increasing in t. Thus, the revision point needs to be as early as possible. On the other hand, if cost values are decreasing in each stage, then the value of the adaptive approach has the same value in all time periods as $\bar{a}^-(t)$ is same for every $2 \leq t \leq T$. Thus, revision time does not affect the analytical bounds for this setting. We can generalize this result as follows by letting the maximum cost value over the scenario tree as $\hat{t}^C = \min\{t \in \{1, \cdots, T\} : a_n = a^*, n \in S_t\}$.

Proposition 11. *Under a general cost structure for a_n with demand values $\{\delta_n\}_{n\in\mathcal{T}} \approx \delta$, $\bar{a}^-(t)$ is monotonically nondecreasing until the period $\hat{t}^C$, and it remains constant afterwards.*

The proof of Proposition 11 is immediate by following the definition of $\bar{a}^-(t)$. This proposition shows that if the maximum cost value is in the root node of the scenario tree, then the value of the adaptive approach is not affected by the revision decision. Otherwise, the revision point needs to be selected as early as possible before observing the highest cost value of the scenario tree.

VATS for the Capacity Expansion Planning Problem

In this section, we extend our results on the single-resource subproblem (4.18) to the capacity expansion planning problem (4.13) under a given revision decision for each resource. We derive analytical bounds on the objective of the adaptive two-stage approach in comparison to two-stage and multi-stage stochastic models. To derive the desired bounds, we utilize the linear programming relaxations of the capacity expansion planning problem.

Proposition 12. *Let $\{y_n^{TLP}\}_{n\in\mathcal{T}}$ be the optimal capacity allocation decisions to the linear programming relaxation of the two-stage version of the stochastic capacity expansion planning problem* (4.13). *Then,*

$$V^{TS} - V^{ATS}(t^*) \geq \sum_{i=1}^{I} (v_i^T - v_i^R(t_i^*) - \max_{n\in\hat{\mathcal{T}}_i(t_i^*)} \{\lceil\delta_{in}\rceil - \delta_{in}\} a_{i1}), \tag{4.25}$$

where v_i^T and $v_i^R(t_i^*)$ are the optimal objective values of the linear programming relaxations of the models (4.17) *and* (4.18) *under $\delta_{in} = [A_n y_n^{TLP}]_i$.*

Proof. We first observe that $V^{TS} \geq \sum_{n\in\mathcal{T}} p_n b_n^\top y_n^{TLP} + \sum_{i=1}^{I} v_i^T$. Next, we note that $\{y_n^{TLP}\}_{n\in\mathcal{T}}$ is a feasible solution to the adaptive two-stage problem under t^*. Thus, $V^{ATS}(t^*) \leq \sum_{n\in\mathcal{T}} p_n b_n^\top y_n^{TLP} + \sum_{i=1}^{I} o_i^R(t_i^*)$, where

$$o_i^R(t_i^*) = \min\{ \sum_{n\in\mathcal{T}_i(t_i^*)} \hat{p}_n \hat{a}_{in} x_{in} : \quad \sum_{m\in P(n)} x_{im} \geq \hat{\delta}_{in}, x_{in} \in \mathbb{Z}_+ \; \forall n \in \mathcal{T}_i(t_i^*)\}.$$

Additionally, we can represent $v_i^R(t_i^*)$ as follows

$$\begin{aligned} v_i^R(t_i^*) &= \min\{ \sum_{n\in\mathcal{T}_i(t_i^*)} \hat{p}_n \hat{a}_{in} x_{in} : \quad \sum_{m\in P(n)} x_{im} \geq \hat{\delta}_{in}, x_{in} \in \mathbb{R}_+ \; \forall n \in \mathcal{T}_i(t_i^*)\} \\ &= \max\{ \sum_{n\in\mathcal{T}_i(t_i^*)} \hat{\delta}_{in} \pi_{in} : \quad \sum_{m\in\mathcal{T}_i(t_i^*)} \pi_{im} \leq \hat{p}_n \hat{a}_{in}, \pi_{in} \in \mathbb{R}_+ \; \forall n \in \mathcal{T}_i(t_i^*)\}. \end{aligned}$$

Using Proposition 6 and linear programming duality, we can reexpress $o_i^R(t_i^*)$ as

$$\begin{aligned} o_i^R(t_i^*) &= \min\{ \sum_{n\in\mathcal{T}_i(t_i^*)} \hat{p}_n \hat{a}_{in} x_{in} : \sum_{m\in P(n)} x_{im} \geq \lceil\hat{\delta}_{in}\rceil, x_{in} \in \mathbb{R}_+ \; \forall n \in \mathcal{T}_i(t_i^*)\} \\ &= \max\{ \sum_{n\in\mathcal{T}_i(t_i^*)} (\hat{\delta}_{in} + (\lceil\hat{\delta}_{in}\rceil - \hat{\delta}_{in}))\pi_{in} : \sum_{m\in\mathcal{T}_i(t_i^*)} \pi_{im} \leq \hat{p}_n \hat{a}_{in}, \pi_{in} \in \mathbb{R}_+ \; \forall n \in \mathcal{T}_i(t_i^*)\} \\ &\leq v_i^R(t_i^*) + \max_{n\in\hat{\mathcal{T}}_i(t_i^*)} \{\lceil\hat{\delta}_{in}\rceil - \hat{\delta}_{in}\} \; \max\{ \sum_{n\in\mathcal{T}_i(t_i^*)} \pi_{in} : \\ &\qquad \sum_{m\in\mathcal{T}_i(t_i^*)} \pi_{im} \leq \hat{p}_n \hat{a}_{in}, \pi_{in} \in \mathbb{R}_+ \; \forall n \in \mathcal{T}_i(t_i^*)\} \end{aligned}$$

$$= v_i^R(t_i^*) + \max_{n \in \hat{\mathcal{T}}_i(t_i^*)} \{\lceil \hat{\delta}_{in} \rceil - \hat{\delta}_{in}\} \ \min\{ \sum_{n \in \mathcal{T}_i(t_i^*)} \hat{\hat{p}}_n \hat{a}_{in} x_{in} : \sum_{m \in P(n)} x_{im} \geq 1, x_{in} \in \mathbb{R}_+ \ \forall n \in \mathcal{T}_i(t_i^*)\}$$

$$= v_i^R(t_i^*) + \max_{n \in \hat{\mathcal{T}}_i(t_i^*)} \{\lceil \hat{\delta}_{in} \rceil - \hat{\delta}_{in}\} \ a_{i1}.$$

Here, the last equality follows from the fact that $x_{i1} = 1$ and $x_{in} = 0$ for $n \in \mathcal{T}_i(t_i^*) \setminus \{1\}$ is an optimal solution of the resulting single-resource problem. Combining the above, we demonstrate the desired result. □

Proposition 13. *Let* $\{y_n^{MLP}\}_{n \in \mathcal{T}}$ *be the optimal capacity allocation decisions to the linear programming relaxation of the multi-stage stochastic capacity expansion planning problem* (4.13). *Then,*

$$V^{ATS}(t^*) - V^{MS} \leq \sum_{i=1}^{I} (v_i^R(t_i^*) - v_i^M + \max_{n \in \hat{\mathcal{T}}_i(t_i^*)} \{\lceil \delta_{in} \rceil - \delta_{in}\} a_{i1}), \tag{4.26}$$

where v_i^M *and* $v_i^R(t_i^*)$ *are the optimal objective values of the linear programming relaxations of the models* (4.16) *and* (4.18) *under* $\delta_{in} = [A_n y_n^{MLP}]_i$.

Proof. We first observe that $V^{MS} \geq \sum_{n \in \mathcal{T}} p_n b_n^\top y_n^{MLP} + \sum_{i=1}^I v_i^M$. Next, we have $V^{ATS}(t^*) \leq \sum_{n \in \mathcal{T}} p_n b_n^\top y_n^{MLP} + \sum_{i=1}^I o_i^R(t_i^*)$ as before. Using the same techniques as in the proof of Proposition 12, we obtain the desired result. □

Using the bound in Proposition 13 and the relationship $V^{MS} \leq V^{ATS}(t^*)$ for any revision vector t^*, we can evaluate the performance of the adaptive two-stage approach in comparison to its value under a given t^* vector.

Corollary 2. *Let* V^{ATS} *denote the optimal objective of the adaptive two-stage capacity expansion problem. Then,*

$$V^{ATS}(t^*) - V^{ATS} \leq \sum_{i=1}^{I} (v_i^R(t_i^*) - v_i^M + \max_{n \in \hat{\mathcal{T}}_i(t_i^*)} \{\lceil \delta_{in} \rceil - \delta_{in}\} a_{i1}), \tag{4.27}$$

where $\delta_{in} = [A_n y_n^{MLP}]_i$.

We note that these theoretical results will be used in the solution methodology introduced in the next section for developing heuristic approaches with approximation guarantees.

4.5 Solution Methodology

The adaptive two-stage formulation imposes computational challenges when the revision times of resources are decision variables. In this section, we propose three approximation algorithms to solve the adaptive two-stage equivalent of the capacity expansion planning problem (4.13). Our first two algorithms are based on the bounds derived in Propositions 12 and 13 by selecting the revision points that provide most gain against two-stage and least loss against multi-stage models. In Algorithm 5, we identify the revision point of each resource by maximizing the lower bound of the gain in objective in comparison to two-stage stochastic model. Similarly, in Algorithm 6, we determine the revision points by minimizing the upper bound of the loss in objective in comparison to multi-stage stochastic model.

Algorithm 5 Algorithm Two-stage Relax (**TS-Relax**)

1: Solve the linear programming relaxation of the two-stage version of the stochastic capacity expansion planning problem (4.13) and obtain $\{y_n^{TLP}\}_{n\in\mathcal{T}}$.

2: Compute $\delta_{in} = [A_n y_n^{TLP}]_i$ for all $i = 1, \cdots, I, n \in \mathcal{T}$.

3: **for all** $i = 1, \ldots, I$ **do**

4: Find t_i^* that maximizes the lower bound on $v_i^T - v_i^R(t_i^*) - \max_{n\in\hat{\mathcal{T}}_i(t_i^*)}\{\lceil\delta_{in}\rceil - \delta_{in}\}a_{i1}$ using the bounds (4.20).

5: **end for**

6: Solve the adaptive two-stage version of the stochastic capacity expansion planning problem (4.13) for $\{x_n, y_n\}_{n\in\mathcal{T}}$ given the $\{t_i^*\}_{i=1}^I$ values.

Algorithm 6 Algorithm Multi-stage Relax (**MS-Relax**)

1: Solve the linear programming relaxation of the multi-stage stochastic capacity expansion planning problem (4.13) and obtain $\{y_n^{MLP}\}_{n\in\mathcal{T}}$.

2: Compute $\delta_{in} = [A_n y_n^{MLP}]_i$ for all $i = 1, \cdots, I$, $n \in \mathcal{T}$.

3: **for all** $i = 1, \ldots, I$ **do**

4: Find t_i^* that minimizes the upper bound on $v_i^R(t_i^*) - v_i^M + \max_{n\in\hat{\mathcal{T}}_i(t_i^*)}\{\lceil\delta_{in}\rceil - \delta_{in}\}a_{i1}$ using the bounds (4.21).

5: **end for**

6: Solve the adaptive two-stage version of the stochastic capacity expansion planning problem (4.13) for $\{x_n, y_n\}_{n\in\mathcal{T}}$ given the $\{t_i^*\}_{i=1}^I$ values.

In addition to these, we propose another approximation algorithm in Algorithm 7 by first solving a relaxation of the adaptive two-stage stochastic program to identify the revision points. Then, we obtain the capacity expansion and allocation decisions under the resulting revision decisions.

Algorithm 7 Algorithm Adaptive Two-stage Relax (**ATS-Relax**)

1: Solve a relaxation of the adaptive two-stage version of the stochastic capacity expansion planning problem (4.13) where $\{x_n\}_{n\in\mathcal{T}}$ decisions are continuous, and obtain $\{r_{i,t}^{ALP}\}$ for $i = 1, \cdots, I, t = 1, \cdots, T$.

2: Let $t_i^* = \sum_{t=1}^T t r_{i,t}^{ALP}$ for all $i = 1, \cdots, I$.

3: Solve the adaptive two-stage version of the stochastic capacity expansion planning problem (4.13) for $\{x_n, y_n\}_{n\in\mathcal{T}}$ given the $\{t_i^*\}_{i=1}^I$ values.

We note that Corollary 2 provides an upper bound to compare the objectives of the true adaptive two-stage program with the adaptive program under a given revision point. Using this result, we can demonstrate the approximation guarantees of Algorithms 5 - 7 according to their choices of revision decisions. We can further improve this bound for Algorithm 7 as follows.

Proposition 14. *Let $V^{ATS-Relax}$ and $t^{ATS-Relax}$ denote the objective and the revision vector of the adaptive two-stage program under the solutions found in Algorithm 7. Then,*

$$V^{ATS-Relax} - V^{ATS} \leq \sum_{i=1}^{I} \Big(\max_{n \in \hat{\mathcal{T}}_i(t_i^{ATS-Relax})} \{(\lceil \delta_{in} \rceil - \delta_{in})\} a_{i1} \Big), \tag{4.28}$$

where $\delta_{in} = [A_n y_n^{LP}]_i$ and y^{LP} is the capacity allocation decisions found in Step 1 of Algorithm 7.

Proof. We denote the objective value of the adaptive two-stage version of the stochastic capacity expansion planning problem (4.13) under a given (x, y, t) decision as

$$f(x, y, t) := \sum_{i=1}^{I} \sum_{n \in \mathcal{T}_i(t_i)} \hat{p}_n \hat{a}_{in} x_{in} + \sum_{n \in \mathcal{T}} p_n b_n^\top y_n,$$

where $\mathcal{T}_i(t_i)$ corresponds to the compressed tree under the revision decision t_i, and $\hat{p}_n = \sum_{m \in \hat{C}_n} p_m$, $\hat{a}_{in} = \sum_{m \in \hat{C}_n} p_m a_{im}$ as discussed in formulation (4.19).

We represent the solution corresponding to $ATS - Relax$ algorithm as ($x^{ATS-Relax}$, $y^{ATS-Relax}$, $t^{ATS-Relax}$), and the solution of the true adaptive two-stage program as (x^{ATS}, y^{ATS}, t^{ATS}). Consequently, we define the bound between the two approaches as

$$\begin{aligned}
V^{ATS-Relax} - V^{ATS} =& f(x^{ATS-Relax}, y^{ATS-Relax}, t^{ATS-Relax}) - f(x^{ATS}, y^{ATS}, t^{ATS}) \\
\leq& f(x^{ATS-Relax}, y^{ATS-Relax}, t^{ATS-Relax}) - f(x^{LP}, y^{LP}, t^{LP}) \\
=& f(x^{ATS-Relax}, y^{ATS-Relax}, t^{ATS-Relax}) - f(x^{ATS-Relax}, y^{LP}, t^{LP}) \\
&+ f(x^{ATS-Relax}, y^{LP}, t^{LP}) - f(x^{LP}, y^{LP}, t^{LP}) \\
\leq& f(x^{ATS-Relax}, y^{LP}, t^{LP}) - f(x^{LP}, y^{LP}, t^{LP}),
\end{aligned}$$

where (x^{LP}, y^{LP}, t^{LP}) represents the solution of the relaxation of the true adaptive two-stage program when x decisions are relaxed to be continuous. We note that $t^{LP} = t^{ATS-Relax}$

by the definition of Algorithm 7. Next, we can state the following

$$f(x^{ATS-Relax}, y^{LP}, t^{LP}) - f(x^{LP}, y^{LP}, t^{LP}) = \sum_{i=1}^{I} \sum_{n \in \mathcal{T}_i(t_i^{LP})} \hat{p}_n \hat{a}_{in}(x_{in}^{ATS-Relax} - x_{in}^{LP}),$$

Here, $\hat{p}_n$ and $\hat{a}_n$ are computed specifically for the compressed tree under the revision decision t^{LP}. Using the analysis in the proof of Theorem 6 in [79], the above expression reduces to

$$f(x^{ATS-Relax}, y^{LP}, t^{LP}) - f(x^{LP}, y^{LP}, t^{LP}) = \sum_{i=1}^{I} (\max_{n \in \hat{\mathcal{T}}_i(t_i^{ATS-Relax})} \{(\lceil \delta_{in} \rceil - \delta_{in})\} a_{i1}).$$

Combining the above, we obtain the desired result. □

We can obtain a simpler upper bound on the optimality gap of the ATS-Relax Algorithm derived in (4.28). Specifically, we have $V^{ATS-Relax} - V^{ATS} \leq \sum_{i=1}^{I} a_{i1}$ as $\lceil \delta_{in} \rceil - \delta_{in} \leq 1$ for all resources i and nodes n. We note that this optimality gap is irrespective of the number of stages, scenario tree structure and problem data. The gap only depends on the capacity acquisition cost of each resource at the first stage. Combining the above, we derive the asymptotic convergence of the ATS-Relax Algorithm as follows.

Corollary 3. *The Algorithm 7 asymptotically converges to the true adaptive program, i.e.*

$$\lim_{T \to \infty} \frac{V^{ATS-Relax} - V^{ATS}}{T} = 0.$$

4.6 Computational Analysis

We illustrate our results on an application to generation capacity expansion planning, which contains the problem structure studied in Section 4.4. In Section 4.6.1, we present the experimental setup and the details of the problem formulation. In Section 4.6.2, we provide an extensive computational study by demonstrating the value of the adaptive two-stage approach and comparing the performances of the different solution algorithms on various scenario tree structures. We also present a detailed optimal solution for a specific generation expansion problem and provide insights regarding the proposed approach.

4.6.1 Experimental Setup and Optimization Model

Generation capacity expansion planning is a well-studied problem in literature to determine the acquisition and capacity allocation decisions of different types of generation resources over a long-term planning horizon. Our aim is to optimize the yearly investment decisions of different generation resources while producing energy within the available capacity of each generation resource and satisfying the overall system demand in each subperiod. We consider four types of subperiods within a year, namely peak, shoulder, off-peak and base, depending on their demand level. We assume a restriction on the number of generation resources to purchase until the end of the planning horizon. We consider six different types of generation resources for investment decisions, namely nuclear, coal, natural gas-combined cycle (NG-CC), natural gas-gas turbine (NG-GT), wind and solar. We utilize the data set presented in [80], which is based on a technical report of U.S. Department of Energy [81]. This data set is used for computing capacity amounts and costs associated with each type of generation resource.

Using the predictions in [82], we consider a 10% reduction in the capacity acquisition costs of the renewable generation resources in each year. Additionally, we assume 15% and 10% yearly increases in total for fuel prices and operating costs of natural gas and coal type generation resources, respectively. For other types of generators, we take into account 3% increase in fuel prices [80]. To promote the renewable resources, we allow at most 20% capacity expansion for the traditional generation types in terms of their initial capacities. We assume that the generators are available for production in the period that the acquisition decision is made, similar to [83, 67].

The source of uncertainty of our model is the demand level of each subperiod in set $\mathcal{K}$ throughout the planning horizon. We adopt the scenario generation procedure presented in [62] for constructing the scenario tree. At the beginning of the planning horizon, we start with an initial demand level for each subperiod, using the values in [80]. Then, we randomly generate a demand increase multiplier for each node, except the root node, to estimate that node's demand value based on its ancestor node's demand. The details of the scenario tree generation algorithm is presented in Algorithm 8.

Algorithm 8 Scenario Tree Generation with M branches

Obtain the demand values at the root node as $\{d_{k0}\}_{k\in\mathcal{K}}$.
for all $t = 2, \cdots, T$ **do**
    Let $\underline{\alpha}_t \leq \overline{\alpha}_t$ be the demand increase multiplier bounds.
    Find the equisized multiplier interval values as $\{\alpha_t^j\}_{j=0}^{M}$ where $\alpha_t^j = \underline{\alpha}_t + j(\overline{\alpha}_t - \underline{\alpha}_t)/M$.
    for all $k \in \mathcal{K}$ **do**
        for all node $n \in S_t$ **do**
            Find the order of the node in the stage as $j = n \, mod \, M$.
            Generate the demand increase multiplier $\beta_t^j \sim U(\alpha_t^j, \alpha_t^{j+1})$.
            Define $d_{kn} := \beta_t^j d_{ka(n)}$.
        end for
    end for
end for

Following the studies [62, 67], we represent the generation capacity expansion planning problem as a stochastic program. We reformulate this problem as an adaptive two-stage program, where the capacity acquisition decisions can be revised once within the planning horizon. The sets, decision variables, and parameters are defined as follows:

Sets:

$\mathcal{I}$	Set of generation types for expansion
$\mathcal{K}$	Set of subperiod types within each period

Parameters:

T	Number of periods
n_i^0	Number of generation resource type i at the beginning of planning
n_i^{max}	Maximum number of generation resource type i at the end of planning
m_i^{max}	Maximum capacity of a type i generation resource in MWh
$m_i'^{max}$	Maximum effective capacity of a type i generation resource in MWh
l_i	Peak contribution ratio of a type i generation resource
c_{in}	Capacity acquisition cost per MWh of a generation resource type i at node n
f_{in}	Fix operation and maintenance cost per MWh of a generation resource type i at node n
g_{in}	Fuel price per MWh of a generation resource type i at node n
k_{in}	Generation cost per MWh of a generation resource type i at node n
h_{kt}	Number of hours of subperiod k in period t
d_{kn}	Hourly demand at subperiod k at node n
w	Penalty per MWh for demand curtailment
r	Yearly interest rate

Decision variables:

x_{in}	Number of generation resource type i acquisition at node n
u_{ikn}	Hourly generation amount of generator type i at subperiod k at node n
v_{kn}	Demand curtailment amount at subperiod k at node n
t_i^*	Revision point for acquisition decisions of generation resource type i

We then formulate the generation expansion planning problem as follows:

$$\min_{x,u,v,t^*} \quad \sum_{n\in\mathcal{T}}\sum_{i\in\mathcal{I}} p_n \frac{1}{(1+r)^{t_n-1}} \left(\left(c_{in} + \sum_{t=t_n}^{T} \frac{f_{in}}{(1+r)^{t-t_n}} \right) m_i^{max} x_{in} + \sum_{k\in\mathcal{K}} (g_{in}+k_{in}) h_{kt_n} u_{ikn} \right) + \sum_{n\in\mathcal{T}}\sum_{k\in\mathcal{K}} p_n \frac{1}{(1+r)^{t_n-1}} w h_{kt_n} v_{kn} \tag{4.29a}$$

$$\text{s.t.} \quad u_{ikn}\frac{1}{m_i'^{max}} \leq n_i^0 + \sum_{m\in P(n)} x_{im} \quad \forall n \in \mathcal{T} \tag{4.29b}$$

$$\sum_{i\in I} l_i u_{ikn} + v_{kn} \geq d_{ns} \quad \forall n \in \mathcal{T}, \quad \forall k \in \mathcal{K} \tag{4.29c}$$

$$\sum_{m\in P(n)} x_{im} \leq n_i^{max} \quad \forall n \in S_T, \quad \forall i \in \mathcal{I} \tag{4.29d}$$

$$x_{im} = x_{in} \quad \forall m,n \in S_t, \quad t < t_i^*, \quad \forall i \in \mathcal{I} \tag{4.29e}$$

$$x_{im} = x_{in} \quad \forall m,n \in S_t \cap \mathcal{T}(j), \quad j \in S_{t_i^*}, \quad t \geq t_i^*, \quad \forall i \in \mathcal{I} \tag{4.29f}$$

$$x_{in} \in \mathbb{Z}_+, t_i^* \in \{1,\cdots,T\}, u_{ikn}, v_{kn} \in \mathbb{R}_+ \quad \forall n \in \mathcal{T}. \tag{4.29g}$$

The objective function (4.29a) aims to minimize the expected costs of capacity acquisition, allocation and demand curtailment. In order to compute the capacity acquisition costs, we consider the building cost of the generation unit in addition to its maintenance and operating costs for the upcoming periods. For computing the cost of capacity allocation decisions, we use fuel prices and production cost associated with each type of generation resource. All of the costs are then discounted to the beginning of the planning horizon. Constraint (4.29b) ensures that the production amount in each subperiod is restricted by the total available capacity for each type of generation resource. Constraint (4.29c) guarentees that system demand is satisfied, and constraint (4.29d) restricts the acquisition decision throughout the planning horizon. The provided formulation (4.29) is a special form of the capacity expansion problem studied (4.14). In particular, constraints (4.29b) and (4.29d) correspond to constraint (4.13b), and constraint (4.29c) refers to constraint (4.13c). We note that constraint (4.29d) becomes redundant in the linear programming relaxation of the subproblem (4.15) due to Proposition 2 in [67].

Constraints (4.29e) and (4.29f) represent the adaptive two-stage relationship for the capacity acquisition decisions such that the acquisition decision of each resource type i can be revised at t_i^*. We note that the presented formulation is nonlinear, however it can be reformulated as a mixed-integer linear program by defining additional variables for the revision decisions, as shown in (4.11). In our computational experiments, we solve the resulting mixed-integer linear program.

4.6.2 Computational results

In this section, we first demonstrate the value of the adaptive two-stage approach in comparison to two-stage stochastic programming under different scenario tree structures and demand characteristics. Secondly, we examine the performances of the different algorithms introduced in Section 4.5 for solving the adaptive two-stage problem. Finally, we analyze a generation capacity expansion plan under the adaptive two-stage approach to discuss its practical implications.

To construct our computational testbed, we generate scenario trees for demand values using Algorithm 8 when number of branches at each period, namely M, is equal to 2 or 3. We examine scenario trees with number of stages T ranging from 3 to 10. Consequently, the number of nodes in the scenario trees are in the range of 7 to 29524, corresponding to $(M, T) = (2, 3)$, and $(M, T) = (3, 10)$, respectively. We solve the optimization problem (4.29) under five different randomly generated scenario trees for each (M, T) pair. We report the average of these replications to represent the performances of the proposed methods under various instances. We conduct our experiments on an Intel i5 2.20 GHz machine with 8 GB RAM. We implement the algorithms in Python using Gurobi 7.5.2 with a relative optimality gap of 0.1% and time limit of 2 hours.

Value of Adaptive Two-Stage Approach

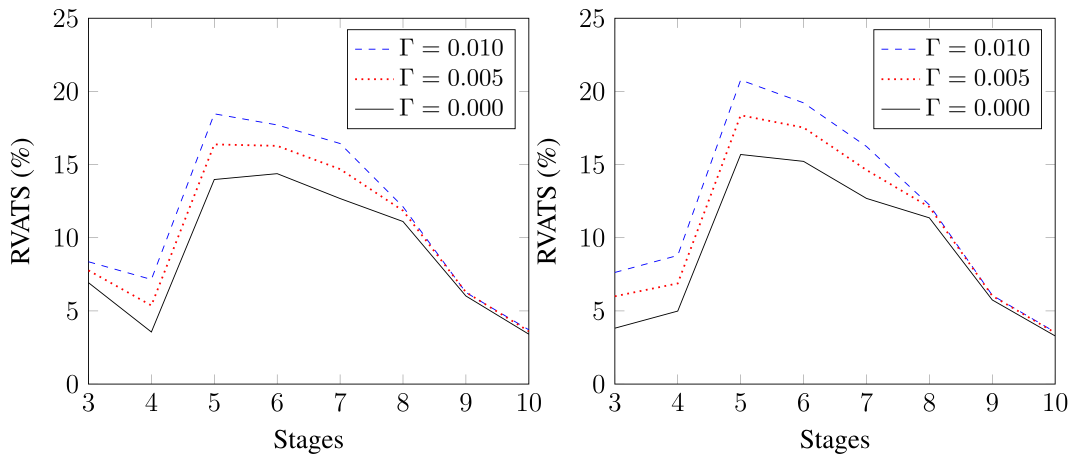
To demonstrate the performance of the adaptive two-stage approach in comparison to two-stage stochastic programming, we define the relative value of adaptive two-stage approach (RVATS) by extending the definition of VATS. Specifically, we let

$$\text{RVATS } (\%) = 100 \times \frac{(V^{TS} - V^{ATS})}{V^{TS}}, \tag{4.30}$$

where V^{TS} and V^{ATS} are the objective values corresponding to the two-stage model and adaptive two-stage model, respectively.

We illustrate the RVATS under various scenario tree structures in Figure 4.5. Figures 4.5a and 4.5b show the behavior of the adaptive two-stage approach under scenario trees with 2 and 3 branches, each with three different demand patterns, respectively. For constructing the demand patterns, we examine the cases with an increasing variance under a constant mean for demand multipliers. Specifically, we consider the demand multipliers in Algorithm 8 as $\underline{\alpha}_t = 1.00 - \Gamma t$ and $\overline{\alpha}_t = 1.20 + \Gamma t$ for all stages $t = 2, \cdots, T$ where $\Gamma = 0, 0.005, 0.01$. To compute RVATS, we utilize V^{ATS} for all stages in 2-branch trees and until stage 7 in 3-branch trees. For larger trees for 3-branch case, we report a lower bound on the RVATS due to the computational difficulty of solving ATS to optimality, as discussed in Section 4.4. In particular, we replace V^{ATS} in Equation 4.30 with $V^{TS-Relax}$, where $V^{TS-Relax}$ represents the objective value of the adaptive two-stage model under the solution of the Algorithm TS-Relax.

We observe RVATS between 3-19% over 2-branch, and 3-21% over 3-branch scenario trees, demonstrating the significant gain of revising decisions over static policies. The largest gain is obtained when the number of stages is equal to 5 for both cases. As the number of stages increases, the adaptive two-stage approach becomes more and more similar to the two-stage stochastic programming. Consequently, the relative gain of the proposed approach decreases for the problems with longer planning horizons.



(a) 2-branch Results

(b) 3-branch Results

Figure 4.5: Value of Adaptive Two-Stage on Instances with Different Variability

We note that the scenario tree has more variability in each stage when the parameter Γ increases. We observe higher gain of the adaptive two-stage approach for the scenario trees with larger variability levels. Additionally, 3-branch scenario trees have higher RVATS values than 2-branch trees due to the larger variance of demand values in each stage. These results demonstrate an even better performance of the adaptive two-stage approach in comparison to two-stage when the scenario tree has larger variability.

Performance of Solution Algorithms

In this section, we examine the performance of the solution algorithms for the capacity expansion planning problem from different perspectives. Tables 4.2 and 4.3 provide the computational results of the solution algorithms for 2-branch and 3-branch scenario trees, respectively. The computational experiments are under the default demand multiplier setting for the scenario trees when $\Gamma = 0.005$, i.e. $\underline{\alpha}_t = 1.00 - 0.005t$ and $\overline{\alpha}_t = 1.20 + 0.005t$ for all stages $t = 2, \cdots, T$. The column "Time" represents the solution time in terms of seconds, and the column "Gain" corresponds to the percentage gain of the studied method in comparison to two-stage stochastic model. We note that we repeat the experiments with larger values of Γ and obtain consistent results.

The algorithms MS-Relax and ATS-Relax provide optimality gaps for the adaptive two-stage program. In particular, these algorithms obtain a lower bound to the adaptive two-stage program by solving relaxations of the multi-stage model and adaptive two-stage model in Step 1 of the Algorithm 5 and Algorithm 7. Since both algorithms construct a feasible solution, they provide an upper bound for the desired problem. Combining the lower and upper bounds, we construct optimality gap for the adaptive two-stage approach. The gap values are reported in the column "Gap" in Tables 4.2 and 4.3. These values provide performance metrics for the solution algorithms in approximating the adaptive two-stage model.

We first note that the ATS approach provides the most gain against the two-stage approach, as expected. Among the approximation algorithms, Algorithm ATS-Relax has the highest gain despite its computational disadvantage. The gain of Algorithm ATS-Relax is very close to the gain of ATS demonstrating the success of the proposed algorithm in approximating the adaptive two-stage approach. Our computational results also highlight the asymptotic convergence of Algorithm ATS-Relax as shown in Corollary 3. Specifically, as number of stages increases, the optimality gap provided by Algorithm ATS-Relax becomes closer to zero. However, we observe significant computational benefits of adopting Algorithms TS-Relax and MS-Relax as they provide notable speedups in comparison to ATS. We also observe that the computational complexity of the ATS and ATS-Relax approaches are more sensitive to the instance size in comparison to Algorithms TS-Relax and MS-Relax.

Table 4.2: Performance of Solution Algorithms on 2-branch Results

	ATS		TS-Relax		MS-Relax			ATS-Relax		
Stage	Time(s)	%Gain	Time(s)	%Gain	Time(s)	%Gain	%Gap	Time(s)	%Gain	%Gap
3	0.16	7.77	0.16	6.01	0.17	4.77	5.47	0.24	7.56	1.25
4	0.58	5.37	0.22	0.79	0.27	4.95	2.80	0.56	5.31	1.18
5	1.42	16.38	0.31	10.70	0.30	14.80	3.76	1.19	16.36	0.06
6	15.34	16.28	0.61	11.41	0.57	13.74	7.84	5.17	16.26	0.04
7	34.05	14.69	1.19	13.00	1.17	11.91	9.56	16.49	14.69	0.02
8	103.25	11.85	3.15	11.22	2.21	9.65	8.48	35.19	11.84	0.03
9	234.10	6.28	5.14	5.79	4.42	4.86	5.46	94.01	6.27	0.02
10	930.34	3.61	12.37	3.10	9.76	2.91	3.09	370.83	3.59	0.03

Table 4.3: Performance of Solution Algorithms on 3-branch Results

	ATS		TS-Relax		MS-Relax			ATS-Relax		
Stage	Time(s)	%Gain	Time(s)	%Gain	Time(s)	%Gain	%Gap	Time(s)	%Gain	%Gap
3	0.48	6.00	0.30	5.32	0.32	4.48	3.85	0.71	5.37	1.57
4	1.52	6.88	0.55	1.20	0.50	6.31	2.67	1.70	6.72	1.23
5	28.30	18.37	1.07	13.73	0.99	16.77	3.90	8.91	18.34	0.10
6	162.22	17.53	3.01	11.04	3.02	11.22	12.06	40.94	17.51	0.02
7	721.34	14.63	13.53	14.57	10.21	11.63	11.64	331.39	14.61	0.02
8	7200.00	-	73.17	12.10	77.97	8.80	10.48	7200.00	-	-
9	7200.00	-	447.11	6.03	265.36	4.14	6.65	7200.00	-	-
10	7200.00	-	2409.83	3.50	1410.35	2.59	3.62	7200.00	-	-

The size of the scenario tree and consequently the problem size increase as the number of stages and the number of branches get larger. We observe that the computational complexity of the ATS and ATS-Relax approaches are more sensitive to the instance size in comparison to Algorithms TS-Relax and MS-Relax. In terms of gain against two-stage approach, we examine higher gains in 3-branch scenario trees demonstrating the effect of higher variance on the performance of the proposed approach.

Discussion on Capacity Expansion Plans

In this section, we examine a particular instance to analyze the generation capacity expansion plan of the adaptive two-stage model and compare it with that of the two-stage model. Figure 4.6 illustrates a generation expansion plan under the adaptive two-stage model over a five-year planning horizon. For each node of the tree, we show the acquired maximum effective capacity amount of each generation resource type $i \in \{1, \cdots, 6\}$ in the order of nuclear, coal, NG-CC, NG-GT, wind and solar. The revision times of each resource type are determined by the model as 2, 3, 5, 5, 4, 5, respectively, and the expansion amounts of those periods are denoted in bold. We report the capacity expansion decisions of a certain resource type i in node n if period $t_n = t_i^*$ or $x_{in} > 0$. For notation simplicity, we omit the node index in the illustration. Figure 4.7 shows the generation expansion plan of the same instance under two-stage model. We note that the initial effective capacities of the resources are 4860, 2618, 2004, 1883, 134, 131 MWh respectively.

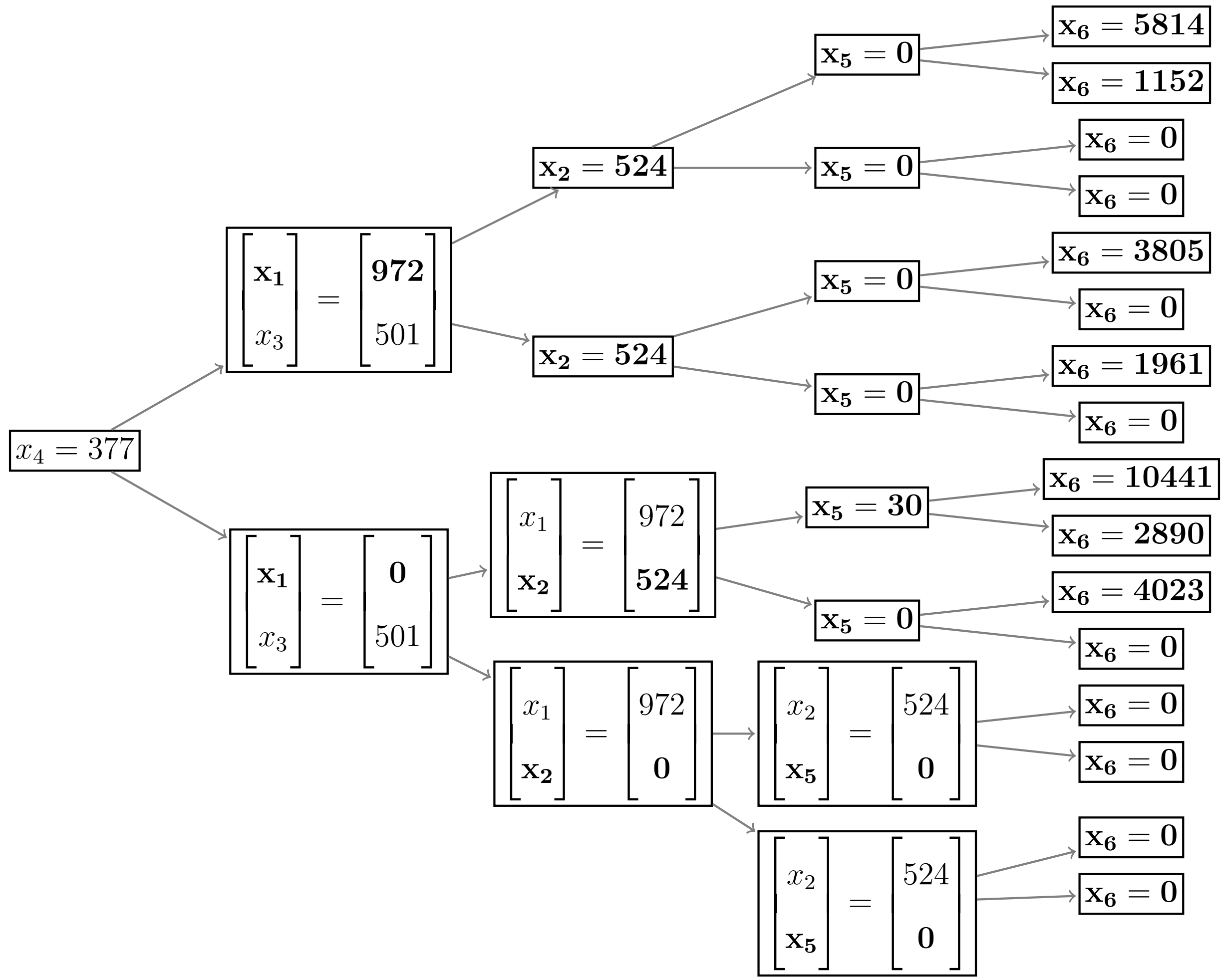


Figure 4.6: Generation expansion plan under adaptive two-stage model.

$$\boxed{x_4 = 377} \longrightarrow \boxed{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 972 \\ 524 \\ 501 \end{bmatrix}} \longrightarrow \boxed{x = 0} \longrightarrow \boxed{x_5 = 1452} \longrightarrow \boxed{x_6 = 509}$$

Figure 4.7: Generation expansion plan under two-stage model.

We observe that the adaptive two-stage approach provides more flexibility compared to the two-stage by allowing an update for each generation type's expansion plan. Since the two-stage approach is less adaptive to the uncertainty, it brings forward the expansion times of the resources to satisfy the overall demand of each stage. We note that the expected total capacity expansion and allocation cost of the adaptive two-stage model is \$47.3 billion and the two-stage model is \$56.2 billion, resulting in 15.79% relative gain of the adaptive two-stage approach.

As the investment costs of the renewable resources decrease over time, both of the models tend to expand wind and solar generation capacities at later times in the planning. Due to this decrease in investment costs and increasing variability of demand in the scenario tree, the revision time of wind and solar capacities are in 4^{th} and 5^{th} periods to adjust the expansion decisions based on the underlying demand. The fuel prices and operating costs associated with traditional generation types increase over time. Thus, the capacities of these types of resources are expanded in the early periods of the planning horizon. To adapt to the demand uncertainty, we observe revisions of expansion decisions at 2^{nd} and 3^{rd} periods for nuclear and coal generation. Since natural gas type generation resources are only expanded in first and second periods, their revision times do not necessarily have a practical implication.

4.7 Concluding Remarks

In this chapter, we propose a stochastic programming approach that determines the best time to revise decisions for problems with partial adaptability. We first present a generic formulation for the proposed adaptive two-stage approach, and demonstrate the importance of the revision times by deriving and illustrating adaptive policies for the multi-period newsvendor problem. We then develop a mixed-integer linear programming reformulation for the generic approach through scenario trees, and show that solving the resulting adaptive two-stage model is NP-Hard. Then, we focus our analyses on a specific problem structure that includes the capacity expansion planning problem under uncertainty. We derive the value of the proposed approach in comparison to two-stage and multi-stage stochastic programming models in terms of the revision times and problem parameters. We also provide a sensitivity analysis on these relative values with respect to the cost and demand parameters of the problem structure studied. We propose solution algorithms that selects the revision times by minimizing the objective gap of the adaptive two-stage approach in comparison to the multi-stage and maximizing the corresponding gap with the two-stage models by benefiting from our analytical analyses. We present an additional solution algorithm based on the relaxation of the adaptive two-stage model, for which we develop an approximation guarantee. In order to illustrate our results, we study a generation capacity expansion planning problem over a multi-period planning horizon. Our extensive computational study highlights up to 21% gain of the adaptive two-stage approach in objective in comparison to two stage model, demonstrating the importance of optimizing revision decisions. We show that this relative gain even increases in scenario trees with higher variability, and provide a computational study illustrating the performance and convergence of the proposed solution algorithms. Finally, we present the practical implications of utilizing the adaptive two-stage approach by examining sample generation capacity expansion plans.

Appendices

APPENDIX A

SUPPLEMENTAL DATA

A.1 Illustrative Instance

In this section, we provide the details of the instance studied in Section 4.4.2. The cost parameter $a_n = 1$ for all $n \in \mathcal{T}$. The demand parameter $\{\delta_n\}_{n \in \mathcal{T}}$ is randomly generated from the distribution $N(\mu, \sigma^2)$, where μ = 30 and σ = 5. We consider a scenario tree with 5 stages, and the resulting values for $\{\delta_n\}_{n \in \mathcal{T}}$ are presented in Figure A.1.

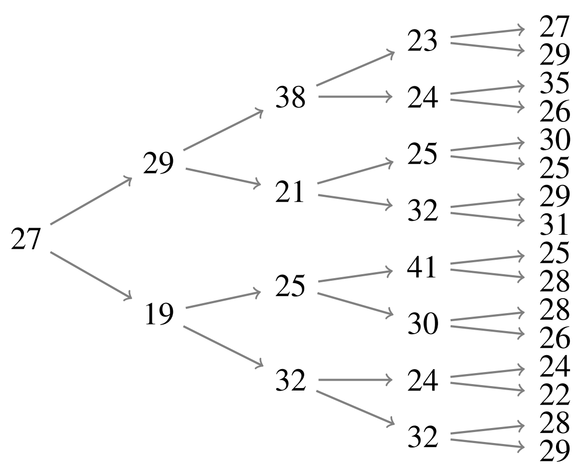


Figure A.1: Demand values for the illustrative instance.

REFERENCES

[1] B. Basciftci, S. Ahmed, N. Z. Gebraeel, and M. Yildirim, "Stochastic optimization of maintenance and operations schedules under unexpected failures," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6755–6765, 2018.

[2] B. Basciftci, S. Ahmed, and N. Z. Gebraeel, *Data-driven maintenance and operations scheduling in power systems under decision-dependent uncertainty*, `http://www.optimization-online.org/DB_HTML/2018/10/6885.html`, 2018.

[3] B. Basciftci, S. Ahmed, and N. Z. Gebraeel, *Adaptive two-stage stochastic programming with an application to capacity expansion planning*, `http://www.optimization-online.org/DB_HTML/2019/06/7244.html`, 2019.

[4] M. Yildirim, N. Z. Gebraeel, and X. A. Sun, "Sensor-driven condition-based generator maintenance scheduling;part i: Maintenance problem," *IEEE Transactions on Power Systems*, vol. 31, no. 6, pp. 4253–4262, 2016.

[5] M. Yildirim, N. Z. Gebraeel, and X. A. Sun, "Sensor-driven condition-based generator maintenance scheduling;part ii: Incorporating operations," *IEEE Transactions on Power Systems*, vol. 31, no. 6, pp. 4263–4271, 2016.

[6] K.-J. Wang and P. H. Nguyen, "Capacity planning with technology replacement by stochastic dynamic programming," *European Journal of Operational Research*, vol. 260, no. 2, pp. 739 –750, 2017.

[7] A. Froger, M. Gendreau, J. E. Mendoza, E. Pinson, and L. Rousseau, "Maintenance scheduling in the electricity industry: A literature review," *European Journal of Operational Research*, vol. 251, no. 3, pp. 695 –706, 2016.

[8] R. Billinton and R. Mo, "Composite system maintenance coordination in a deregulated environment," *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 485–492, 2005.

[9] A. Abiri-Jahromi, M. Fotuhi-Firuzabad, and M. Parvania, "Optimized midterm preventive maintenance outage scheduling of thermal generating units," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1354–1365, 2012.

[10] U. E. Ekpenyong, J. Zhang, and X. Xia, "An improved robust model for generator maintenance scheduling," *Electric Power Systems Research*, vol. 92, pp. 29 –36, 2012.

[11] H. Pandzic, A. J. Conejo, and I. Kuzle, "An epec approach to the yearly maintenance scheduling of generating units," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 922–930, 2013.

[12] A. J. Conejo, R. Garcia-Bertrand, and M. Diaz-Salazar, "Generation maintenance scheduling in restructured power systems," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 984–992, 2005.

[13] E. L. Silva, M. Morozowski, L. G. S. Fonseca, G. C. Oliveira, A. C. G. Melo, and J. C. O. Mello, "Transmission constrained maintenance scheduling of generating units: A stochastic programming approach," *IEEE Transactions on Power Systems*, vol. 10, no. 2, pp. 695–701, 1995.

[14] M. K. C. Marwali and S. M. Shahidehpour, "Integrated generation and transmission maintenance scheduling with network constraints," *IEEE Transactions on Power Systems*, vol. 13, no. 3, pp. 1063–1068, 1998.

[15] Y. Fu, M. Shahidehpour, and Z. Li, "Security-constrained optimal coordination of generation and transmission maintenance outage scheduling," *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 1302–1313, 2007.

[16] Q. P. Zheng, J. Wang, and A. L. Liu, "Stochastic optimization for unit commitment; a review," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 1913–1924, 2015.

[17] S. P. Canto, "Application of benders decomposition to power plant preventive maintenance scheduling," *European Journal of Operational Research*, vol. 184, no. 2, pp. 759 –777, 2008.

[18] L. Wu, M. Shahidehpour, and T. Li, "Genco's risk-based maintenance outage scheduling," *IEEE Transactions on Power Systems*, vol. 23, no. 1, pp. 127–136, 2008.

[19] L. Wu, M. Shahidehpour, and Y. Fu, "Security-constrained generation and transmission outage scheduling with uncertainties," *IEEE Transactions on Power Systems*, vol. 25, no. 3, pp. 1674–1685, 2010.

[20] N. Gebraeel, M. A. Lawley, R. Li, and J. K. Ryan, "Residual-life distributions from component degradation signals: A bayesian approach," *IIE Transactions*, vol. 37, no. 6, pp. 543–557, 2005.

[21] S. Ahmed, "A scenario decomposition algorithm for 0-1 stochastic programs," *Oper. Res. Lett.*, vol. 41, no. 6, pp. 565–569, 2013.

[22] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, 2011.

[23] T. Athay, R. Podmore, and S. Virmani, “A practical method for the direct analysis of transient stability,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-98, no. 2, pp. 573–584, 1979.

[24] S. Blumsack, “Network topologies and transmission investment under electric-industry restructuring,” PhD Thesis, Carnegie Mellon University, Pittsburgh, Pennsylvania, 2006.

[25] N. Gebraeel, “Sensory-updated residual life distributions for components with exponential degradation patterns,” *Automation Science and Engineering, IEEE Transactions on*, vol. 3, no. 4, pp. 382–393, 2006.

[26] E. M. Stein and R. Shakarchi, *Real Analysis, Measure Theory, Integration, & Hilbert Spaces*. Princeton University Press, 2005, vol. 3.

[27] A. Shapiro, “Monte carlo simulation approach to stochastic programming,” in *Proceedings of the 33Nd Conference on Winter Simulation*, ser. WSC ’01, Arlington, Virginia, 2001, pp. 428–431, ISBN: 0-7803-7309-X.

[28] G. Codato and M. Fischetti, “Combinatorial benders’ cuts for mixed-integer linear programming,” *Operations Research*, vol. 54, no. 4, pp. 756–766, 2006.

[29] Q. Zhai, X. Guan, J. Cheng, and H. Wu, “Fast identification of inactive security constraints in scuc problems,” *IEEE Transactions on Power Systems*, vol. 25, no. 4, pp. 1946–1954, 2010.

[30] B. Hua, Z. Bie, C. Liu, G. Li, and X. Wang, “Eliminating redundant line flow constraints in composite system reliability evaluation,” *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 3490–3498, 2013.

[31] R. Madani, J. Lavaei, and R. Baldick, “Constraint screening for security analysis of power networks,” *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 1828–1838, 2017.

[32] F. Pourahmadi, M. Fotuhi-Firuzabad, and P. Dehghanian, “Application of game theory in reliability-centered maintenance of electric power systems,” *IEEE Transactions on Industry Applications*, vol. 53, no. 2, pp. 936–946, 2017.

[33] G. Levitin, L. Xing, and Y. Dai, “Preventive replacements in real-time standby systems with periodic backups,” *IEEE Transactions on Reliability*, vol. 66, no. 3, pp. 771–782, 2017.

[34] X. S. Si, W. Wang, C. H. Hu, and D. H. Zhou, “Remaining useful life estimation a review on the statistical data driven approaches,” *European Journal of Operational Research*, vol. 213, no. 1, pp. 1 –14, 2011.

[35] D. Kumar and B. Klefsjo, “Proportional hazards model: A review,” *Reliability Engineering & System Safety*, vol. 44, no. 2, pp. 177 –188, 1994.

[36] K. A. Doksum and A. Hyland, “Models for variable-stress accelerated life testing experiments based on wiener processes and the inverse gaussian distribution,” *Technometrics*, vol. 34, no. 1, pp. 74–82, 1992.

[37] N. Gebraeel and J. Pan, “Prognostic degradation models for computing and updating residual life distributions in a time-varying environment,” *IEEE Transactions on Reliability*, vol. 57, no. 4, pp. 539–550, 2008.

[38] L. Bian and N. Z. Gebraeel, “Stochastic methodology for prognostics under continuously varying environmental profiles,” *Statistical Analysis and Data Mining: The ASA Data Science Journal*, vol. 6, no. 3, pp. 260–270, 2013.

[39] H. Liao and Z. Tian, “A framework for predicting the remaining useful life of a single unit under time-varying operating conditions,” *IIE Transactions*, vol. 45, no. 9, pp. 964–980, 2013.

[40] W. van Ackooij, I. Danti Lopez, A. Frangioni, F. Lacalandra, and M. Tahanan, *Large-scale unit commitment under uncertainty: An updated literature survey*, 2018.

[41] M. Yildirim, N. Z. Gebraeel, and X. A. Sun, *Leveraging predictive analytics to control and coordinate operations, asset loading and maintenance*, `http://www.optimization-online.org/DB_HTML/2018/07/6712.html`, 2018.

[42] V. Goel and I. E. Grossmann, “A class of stochastic programs with decision dependent uncertainty,” *Mathematical Programming*, vol. 108, no. 2, pp. 355–394, 2006.

[43] A. H. Elwany and N. Z. Gebraeel, “Sensor-driven prognostic models for equipment replacement and spare parts inventory,” *IIE Transactions*, vol. 40, no. 7, pp. 629–639, 2008.

[44] M. B. Cain, R. P. O’Neill, and A. Castillo, *History of optimal power flow and formulations*, http://www.ferc.gov/industries/electric/indus-act/market-planning/opf-papers/acopf-1-history-formulation-testing.pdf, 2012.

[45] J. P. Vielma and G. L. Nemhauser, “Modeling disjunctive constraints with a logarithmic number of binary variables and constraints,” in *Proceedings of the 13th International Conference on Integer Programming and Combinatorial Optimization*, ser. IPCO’08, Bertinoro, Italy: Springer-Verlag, 2008, pp. 199–213.

[46] J. P. Vielma, S. Ahmed, and G. L. Nemhauser, “Mixed-integer models for nonseparable piecewise-linear optimization: Unifying framework and extensions,” *Operations Research*, vol. 58, no. 2, pp. 303–315, 2010.

[47] H. Wilf, *Combinatorial algorithmsan update*. Society for Industrial and Applied Mathematics, 1989, vol. 55.

[48] E. A. Silver, R. Peterson, and D. F. Pyke, *Inventory Management and Production Scheduling*. John Wiley & Sons, 1988.

[49] J. H. Bookbinder and J.-Y. Tan, "Strategies for the probabilistic lot-sizing problem with service-level constraints," *Management Science*, vol. 34, no. 9, pp. 1096–1108, 1988.

[50] S. Tarim and B. G. Kingsman, "The stochastic dynamic production/inventory lot-sizing problem with service-level constraints," *International Journal of Production Economics*, vol. 88, no. 1, pp. 105 –119, 2004.

[51] U. Ozen, M. K. Dogru, and S. A. Tarim, "Static-dynamic uncertainty strategy for a single-item stochastic inventory control problem," *Omega*, vol. 40, no. 3, pp. 348 –357, 2012.

[52] M. Zhang, S. Kucukyavuz, and S. Goel, "A branch-and-cut method for dynamic decision making under joint chance constraints," *Management Science*, vol. 60, no. 5, pp. 1317–1333, 2014.

[53] H. Tunc, O. A. Kilic, S. A. Tarim, and R. Rossi, "An extended mixed-integer programming formulation and dynamic cut generation approach for the stochastic lot-sizing problem," *INFORMS Journal on Computing*, vol. 30, no. 3, pp. 492–506, 2018.

[54] E. Koca, H. Yaman, and M. Akturk, "Stochastic lot sizing problem with nervousness considerations," *Computers & Operations Research*, vol. 94, pp. 23 –37, 2018.

[55] J. A. Van Mieghem, "Commissioned paper: Capacity management, investment, and hedging: Review and recent developments," *Manufacturing & Service Operations Management*, vol. 5, no. 4, pp. 269–302, 2003.

[56] S. Rajagopalan, M. R. Singh, and T. E. Morton, "Capacity expansion and replacement in growing markets with uncertain technological breakthroughs," *Management Science*, vol. 44, no. 1, pp. 12–30, 1998.

[57] J. T. Lin, T.-L. Chen, and H.-C. Chu, "A stochastic dynamic programming approach for multi-site capacity planning in tft-lcd manufacturing under demand uncertainty," *International Journal of Production Economics*, vol. 148, pp. 21 –36, 2014.

[58] J. M. Swaminathan, "Tool capacity planning for semiconductor fabrication facilities under demand uncertainty," *European Journal of Operational Research*, vol. 120, no. 3, pp. 545 –558, 2000.

[59] M. Riis and K. A. Andersen, "Multiperiod capacity expansion of a telecommunications connection with uncertain demand," *Computers & Operations Research*, vol. 31, no. 9, pp. 1427 –1436, 2004.

[60] S. Ahmed and N. V. Sahinidis, "An approximation scheme for stochastic integer programs arising in capacity expansion," *Operations Research*, vol. 51, no. 3, pp. 461–471, 2003.

[61] S. Ahmed, A. J. King, and G. Parija, "A multi-stage stochastic integer programming approach for capacity expansion under uncertainty," *Journal of Global Optimization*, vol. 26, no. 1, pp. 3–24, 2003.

[62] K. J. Singh, A. B. Philpott, and R. K. Wood, "Dantzig-wolfe decomposition for solving multistage stochastic capacity-planning problems," *Operations Research*, vol. 57, no. 5, pp. 1271–1286, 2009.

[63] M. A. H. Dempster, N. H. Pedrón, A. Medova, J. E. Scottnand, and A. Sembos, "Planning logistics operations in the oil industry," *Journal of the Operational Research Society*, vol. 51, no. 11, pp. 1271–1288, 2000.

[64] J. Balasubramanian and I. E. Grossmann, "Approximation to multistage stochastic optimization in multiperiod batch plant scheduling under demand uncertainty," *Industrial & Engineering Chemistry Research*, vol. 43, no. 14, pp. 3695–3713, 2004.

[65] L. Chen and T. H. de Mello, "Re-solving stochastic programming models for airline revenue management," *Annals of Operations Research*, vol. 177, no. 1, pp. 91–114, 2010.

[66] M. Bodur and J. R. Luedtke, *Two-stage linear decision rules for multi-stage stochastic programming*, `https://arxiv.org/pdf/1701.04102.pdf`, 2018.

[67] J. Zou, S. Ahmed, and X. A. Sun, "Partially adaptive stochastic optimization for electric power generation expansion planning," *INFORMS Journal on Computing*, vol. 30, no. 2, pp. 388–401, 2018.

[68] S. Ahmed, *Multistage stochastic optimization*, `https://www.ima.umn.edu/materials/2015-2016/ND8.1-12.16/25386/mssp.pdf`, 2016.

[69] J. Dupačová, *Contamination for multistage stochastic programs*, `http://edoc.hu-berlin.de/18452/9003`, 2006.

[70] J. Dupačová, M. Bertocchi, and V. Moriggia, "Testing the structure of multistage stochastic programs," *Computational Management Science*, vol. 6, no. 2, pp. 161–185, 2009.

[71] M. A. H. Dempster and G. Thompson, *Dynamic portfolio replication using stochastic programming*. Risk Management: Value at Risk and Beyond. Cambridge Univ. Press, 2002, pp. 100–128.

[72] M. Bertocchi, M. Vittorio, and J. Dupačová, "Horizon and stages in applications of stochastic programming in finance," *Annals of Operations Research*, vol. 142, no. 1, pp. 63–78, 2006.

[73] P. Zipkin, *Foundations of Inventory Management*. McGraw-Hill, 2000.

[74] A. P. Ruszczynski and A. Shapiro, *Stochastic programming*. Elsevier Amsterdam, 2003, vol. 10.

[75] G. C. Pflug, "Scenario tree generation for multiperiod financial optimization by optimal discretization," *Mathematical Programming*, vol. 89, no. 2, pp. 251–271, 2001.

[76] A. Shapiro, "Inference of statistical bounds for multistage stochastic programming problems," *Mathematical Methods of Operations Research*, vol. 58, no. 1, pp. 57–68, 2003.

[77] D. Kuhn, *Generalized Bounds for Convex Multistage Stochastic Programs*. Springer, 2005, vol. 548.

[78] M. Garey and D. Johnson, *Computers and Intractability; A Guide to the Theory of NP-Completeness*. New York, NY, USA: W. H. Freeman & Co., 1990.

[79] K. Huang and S. Ahmed, "The value of multistage stochastic programming in capacity planning under uncertainty," *Operations Research*, vol. 57, no. 4, pp. 893–904, 2009.

[80] D. Min, J. hyun Ryu, and D. G. Choi, "A long-term capacity expansion planning model for an electric power system integrating large-size renewable energy technologies," *Computers & Operations Research*, vol. 96, pp. 244 –255, 2018.

[81] V. Black, "Cost and performance data for power generation technologies prepared for the national renewable energy laboratory," U.S. Department of Energy, National Renewable Energy Laboratory, Tech. Rep., 2012.

[82] Y. Liu, R. Sioshansi, and A. J. Conejo, "Multistage stochastic investment planning with multiscale representation of uncertainties and decisions," *IEEE Transactions on Power Systems*, vol. 33, no. 1, 2018.

[83] S. Jin, S. M. Ryan, J.-P. Watson, and D. L. Woodruff, "Modeling and solving a large-scale generation expansion planning problem under uncertainty," *Energy Systems*, vol. 2, no. 3, pp. 209–242, 2011.